

The long war against flu

That the H5N1 strain of bird flu has not yet caused a pandemic is no cause for complacency. Preparations for the inevitable must be redoubled to mitigate the potential devastation.

Five years after the deadly H5N1 avian influenza virus exploded into a global epidemic in birds, it has infected more than 300 people. Happily, it has not yet evolved into a strain that can transmit easily between humans — an event that would trigger a pandemic that could kill tens of millions. But as long as H5N1 continues to be present in animals, that risk persists. And with so many other flu strains out in the world, all constantly evolving, a flu pandemic is inevitable.

This grim reality has spurred basic research into topics such as the 1918 flu virus, cell-receptor biology and evolutionary dynamics, which are collectively yielding insights into the molecular basis of virulence and how viruses adapt to humans. Researchers have also begun to unravel the often fatal clinical events caused by the virus, such as the massive immune response that is a 'cytokine storm', and cell-culture technology is promising to make vaccines available more quickly. Plans by the Bill & Melinda Gates Foundation, the Wellcome Trust and the Pasteur Institute to roadmap this research should help focus priorities for funding, just as similar work has done for neglected diseases.

But improved control measures, especially for H5N1 itself, and public-health infrastructure are our frontline defences against a pandemic. Unfortunately, the overall control picture is bleak. Thailand, Vietnam and China have notched up successes in curbing outbreaks in birds, which is key to minimizing the chance that the virus can pass to humans. But South Korea had its worst outbreak ever in April, and the disease has become endemic in Indonesia, Bangladesh, Vietnam and Egypt. Eradication now seems impossible, and the task of containing the virus has become chronic and costly.

Many countries have made patchy progress in planning how to mitigate a pandemic once it does break out. True, any such plan can only buy time, by using antiviral drugs and restricting movement, until a vaccine is available for the specific strain that has broken out. The Commentary on page 162 endorses what might be an intriguing adjunct: 'pre-pandemic' vaccines, which would be matched not to

the exact pandemic strain, but to earlier variants. Even if these vaccines were only partly effective, advocates argue, they might confer sufficient protection to prevent death or severe disease. Although this idea is untested, it merits consideration — especially as strain-specific vaccines would be available only several months into the pandemic, and even then would be in very short supply. The World Health Organization is planning to stockpile more than 100 million doses of pre-pandemic vaccines, and some nations, including Japan, are considering the same.

But delivering sufficient perfectly matched pandemic vaccine fast enough to make a difference is the critical issue. One promising approach — equipping vaccines with adjuvants that boost their effect, reducing the amount of antigen needed in each dose — is belatedly getting the attention it deserves. Indeed, research is generating vaccine formulations that need so little antigen that timely doses could, in principle, be provided for everyone on Earth using existing plant capacity.

Rapid delivery will require an unprecedented level of international coordination. Plans should be in place so that when a pandemic strikes clinical trials of the strain-specific vaccine begin — as do the manufacture and distribution of the billions of syringes needed to deliver it. There should also be international mechanisms to ensure that developing countries have access to pandemic and pre-pandemic vaccines at low cost.

Surveillance, control of disease in animals, pandemic planning and vaccines — each requires intense, organized and sustained commitment. Even if H5N1 never evolves into a pandemic strain, it serves as a useful wake-up call, revealing just how much more must be done to be better prepared for the inevitable. ■

"Even if H5N1 never evolves into a pandemic strain, it serves as a useful wake-up call."

An unnecessary battle

Neuroscientists and geneticists don't need to be at loggerheads over the biology of mental disorders.

Mental disorders take a staggering health and economic toll. The World Health Organization has estimated that unipolar depressive disorder alone is one of the leading causes of disability worldwide. Schizophrenia, bipolar disorder, autism and the many other psychiatric disorders only add to the misery. Yet progress in understanding the underlying causes of these conditions seems to be moving at a crawl. Genes are surely involved, but decades of futile hunting have made it painfully

clear that the contribution of any single gene to disease is probably minuscule.

How to find these tenuous connections is a contentious scientific debate, with geneticists and neuroscientists at an apparent impasse (see page 154). These two communities must start working together more constructively if they are to crack this challenging problem and ensure that the millions of dollars now flooding into this field are not misspent.

Many geneticists believe that scanning the entire genomes of a massive number of patients will uncover weak gene candidates by sheer statistical power. Yet some neuroscientists dismiss these studies, questioning the utility of indiscriminately seeking a swath of genes that are weakly associated with a condition, and that are unlikely to have relevance to its biology. Another concern is that diagnoses of psychiatric

disorders are often based on fuzzy, subjective and inconsistent criteria. The neuroscientists feel that such untargeted searches will be futile if the population labelled as having one condition — such as schizophrenia — in fact includes many different clinical disorders.

The antipathy is mutual. Some neuroscientists take what is to a geneticist an unacceptably small number of subjects — sometimes not even patients — and focus on establishing links between variants of a candidate gene and brain anatomy or function. Geneticists are often not persuaded by the statistics and question the focus on a few candidate genes whose link to disease has never been well established. Neither side seems to agree that the outcomes of the other's studies represent constructive leads. Nor is either side willing to compromise its standards. As a result, it is difficult to publish in the intersection between these two fields.

This conflict is unnecessary. Both communities will gain by learning from the other. Teams undertaking expensive genome-wide association studies should consult fully with clinical and research neuroscientists. If a single diagnostic label such as 'schizophrenia' is inadequate, then more detailed clinical information on patients must be collected when

they are recruited. The studies might also benefit from more quantifiable measures of behaviour and neurobiology — such as laboratory assays of anxiety or the functioning of fear circuits in the brain.

And neuroscientists making early forays into genetic associations may find the hard lessons learned by geneticists over the years useful. Geneticists know about the statistical methods and criteria required to draw genetic conclusions that are persuasive and likely to stand the test of time.

The two sides should not lose sight of the fact that they share the same goal: to help patients. In the end, it is likely that both strategies and others will prove valuable in identifying risk factors and developing ways to counteract them. If genome-wide scans are appropriately designed and powered, they should yield the first clues. The neuroscientists' approach may then prove fruitful by showing what biological function is being performed by the genes identified, how that function leads to disease and how it can be altered.

When two fields such as these come together, lowering standards is not an option. Extracting psychiatric genes will require highly rigorous strategies, and on that geneticists and neuroscientists can agree. ■

A social contract

Efforts to inform US military policy with insights from the social sciences could be a win-win approach.

Given the current US administration's notorious lack of respect for science, the efforts of defence secretary Robert Gates merit a mention. Gates took charge of the Pentagon in December 2006 and ever since has been trying to inform the military's thinking with researchers' insights into other cultures and societies. In October, for example, the Department of Defense launched the Human Terrain System (HTS) programme, in which teams of social scientists advise US troops operating in Iraq and Afghanistan. And on 30 June, the defence department signed a memorandum of understanding that will direct some of its money into social and behavioural science research through the National Science Foundation (NSF).

Gates's outreach may owe something to his tenure as president of Texas A&M University in College Station for the four years before his current appointment. But it certainly reflects the hard lessons of Iraq and Afghanistan — where troops with insufficient understanding of the cultural or political landscape have too often exacerbated the insurgency they were trying to control.

With an HTS team, an anthropologist, say, might be present to advise a commander on the subtleties of negotiating with village elders in rural Afghanistan. According to Gates, one commander in Afghanistan says that using an HTS team has cut the number of armed attacks he has had to make by 60%.

The recent NSF deal aims to shape the Pentagon's long-term strategic thinking by funding academic research in areas such as religious fundamentalism, terrorism and cultural change. Gates also hopes that such research could foster entirely new intellectual tools, in much the same way that work during the cold war fostered game theory.

All proposals will be selected for funding through the NSF's standard peer-review process. The research will be unclassified and no restrictions will be placed on researchers' freedom to publish their results — or for that matter, to criticize the defence department.

The two initiatives have received a decidedly mixed response. Some social scientists have enthusiastically embraced the goals of the HTS programme. But the American Anthropological Association (AAA) has formally condemned it, saying that participants would find it difficult or impossible to follow the association's ethical guidelines in a combat zone, and that it could make it more difficult for anthropologists to build trust elsewhere in the world.

The NSF agreement has been widely acclaimed by university administrators, who welcome the extra research money. But it has aroused the suspicion among some researchers that it will distort social science towards military priorities. Of particular concern is the fact that the defence department will have some say in the choice of the NSF's peer reviewers.

So far, Gates and his deputies have tried to be sensitive to these concerns. But continued vigilance is paramount. War is notoriously fraught with ambiguity and moral compromise, and it may well be a temptation for commanders to use the information gathered by HTS researchers to, say, target suspect populations. Such temptations should be resisted wherever possible: in the long run, honest dealings and the do-no-harm principle required by the AAA are in the military's own best interest. So too is avoiding any temptation to load the NSF peer-review panels.

Social scientists, meanwhile, should embrace the opportunities that the AAA pointed out last November in a report on engagement with the military. These include studying military and intelligence organizations from the inside and educating the military about other cultures and societies. Outrage at the current administration should not derail efforts that have potential to be a win-win for all concerned — including, most especially, the people of Iraq, Afghanistan and regions of future conflict. ■

RESEARCH HIGHLIGHTS

A watery carbon bank

Ecosystems **11**, 643–653 (2008)

Dead wood can persist for more than a hundred times longer in rivers and floodplains than on land, sequestering carbon for centuries and even millennia, according to Richard Guyette of the University of Missouri in Columbia and his colleagues.

Although their terrestrial neighbours degrade within decades, the submerged oaks that Guyette's team studied held their carbon for an average of almost 2,000 years. The team found oak wood up to 14,000 years old in northern Missouri streambeds and floodplains.

The samples are among the oldest non-petrified oak trees known in North America. The aged wood retains its rings, and is a potential source of palaeoclimatic data.



R. GUYETTE

NEUROSCIENCE

Rewiring the brain

J. Neurosci. **28**, 6592–6606 (2008)

The brain can recover so well from a stroke that initially paralysed limbs can be moved again. Scientists have discovered how this happens at the level of individual neurons.

Timothy Murphy and Ian Winship of the University of British Columbia in Vancouver induced stroke in adult mice and used an *in vivo* imaging technique called two-photon microscopy to monitor the activity of individual neurons close to the site of damage.

In the first month — when paralysis is usually at its worst — they found that some neurons ditched their speciality for one particular limb and began processing information from multiple limbs. During the following month, as the affected brain region reorganized itself more permanently, those neurons re-specialized to a new single limb.

MATERIALS

Colourful clay

Adv. Mater. doi:10.1002/adma.200702544 (2008)

Nanocrystals made of semiconducting materials glow different colours depending on their size.

Researchers are interested in using them in electronic devices, but first need to figure out how to attach them to more conventional silicon components.

Takeo Ebina and his colleagues at the National Institute of Advanced Industrial Science and Technology in Sendai, Japan, have come up with a technique for doing so. The researchers dissolved nanocrystals made of cadmium

selenide and zinc sulphide in water and mixed it into a clay–polymer mixture. They then poured the clay into moulds and dried it, creating a thin, flexible film filled with evenly spaced crystals. The team was able to make films glow various colours under ultraviolet light (pictured below) and hope the technique will allow nanocrystals to be integrated into optoelectronic devices.

REMOTE SENSING

Rainforest shrinkage

Proc. Natl Acad. Sci. USA **105**, 9439–9444 (2008)

Despite growing international concern about the future of the world's rainforests, the rate of tropical forest clearance has not slowed.

A team used a combination of low- and high-resolution satellite data sets to quantify forest clearing in the humid tropics. Using a probability-based approach blending satellite data retrieved from NASA's Terra and Landsat missions, Ruth DeFries of the University of Maryland in College Park and her colleagues estimate that more than 27 million hectares of rainforest area — roughly 2.4% of the global rainforest cover — were cleared from 2000 to 2005.

The results suggest that tropical forest loss continues at rates roughly similar with those observed in the 1990s.

PLANT BIOLOGY

An egg-speditent defence

Proc. Natl Acad. Sci. USA doi:10.1073/

pnas.0707809105 (2008)

Some plants are known to attract parasitic wasps to eat the eggs laid on them by insects. However, the exact mechanisms of this complex defensive behaviour are not clear.

Nina Fatouros at Wageningen University in the Netherlands and her colleagues have determined that in one case the mechanism is triggered by a specific compound in male insect ejaculate. The ejaculate of male large cabbage white butterflies (*Pieris brassicae*) contains anti-aphrodisiac benzyl cyanide to reduce female re-mating. This compound is then present on or around their eggs.

Application of benzyl cyanide induced chemical changes in the leaves of Brussels sprout plants, Fatouros's team found. As a result, the parasitoid wasp *Trichogramma brassicae* spent more time on treated leaves than on controls. So, for *P. brassicae* males, the anti-aphrodisiac trick has some cost.

ATMOSPHERIC CHEMISTRY

Forgotten gas

Geophys. Res. Lett. **35**, L12810 doi:10.1029/2008GL034542 (2008)

Nitrogen trifluoride (NF₃) has been identified before as a greenhouse gas, but the threat it poses has barely been quantified. Michael Prather and Juno Hsu of the University of California, Irvine, have produced a new estimate of the atmospheric lifetime of NF₃ —



H. TETSUKA

550 years. This is about 25% shorter than the previous value, but, they write, “still far beyond any societal time frames”.

The market for NF_3 has exploded recently, because its initial use in rocket fuel has expanded to include applications in the semiconductor industry. NF_3 is already potentially a bigger contributor to climate warming than the carbon dioxide emissions of some large coal-fired power plants, and global production in 2008 is predicted to be more than 3,500 tonnes — a number that could double by 2010. Yet NF_3 is not included in the greenhouse gases covered by the Kyoto Protocol.

ECOLOGY

Competing keeps bees busy

J. Ecol. doi:10.1111/j.1365-2745.2008.01405.x (2008)
Bumblebees are pollinator generalists, flying to more flower species than most other insects, but a new study suggests that this behaviour depends on competition.

Colin Fontaine and his colleagues at the École Normale Supérieure in Paris recorded the foraging behaviour of individual common bumblebees (*Bombus terrestris*) in an experimental garden where five flowering plant species were available. The researchers varied the number of bees present during the experiment and found that when few bees were present, they visited fewer species of plant.

Pollinator numbers are known to be falling in many regions owing to human disturbance. Worryingly, the resulting reduced competition could lead bumblebees to eschew some plant species.

GEOLOGY

Glacial speeds

Science 321, 111–113 (2008)

The movement of ice sheets increases markedly when meltwater gets involved, according to researchers from Utrecht University in the Netherlands.

Just days after increased meltwater is produced, ice velocity can increase by a factor of four, say Roderik van de Wal and his colleagues. They used 17 years' worth of data from Global Positioning System units drilled into the western edge of Greenland's ice sheet to determine both ice velocity and disappearance rates. Despite detecting larger than previously observed meltwater-induced

bursts in velocity, the team found that, in the long term, ice velocities have actually decreased slightly.

They suggest that ice sheets continually adjust to the amount of meltwater, keeping overall velocities more or less constant.

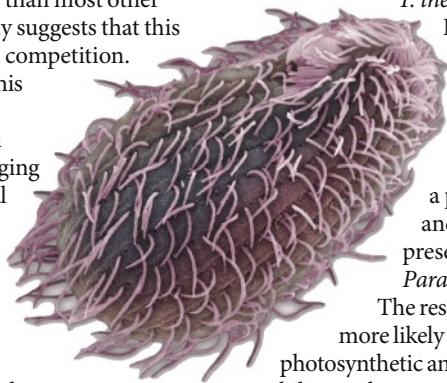
EVOLUTION

Photosynthetic ancestors

Curr. Biol. doi:10.1016/j.cub.2008.05.042 (2008)
Genomic analysis of the hairy single-celled ciliate *Tetrahymena thermophila* (pictured below) suggests that ciliates may once have been photosynthetic. This finding supports the controversial ‘chromalveolate hypothesis’, which proposes a common photosynthetic ancestor for two groups of single-celled eukaryotes: the chromists and the alveolates.

Searching through the tens of thousands of proteins predicted from the genome of *T. thermophila*, an alveolate, Debashish Bhattacharya and his colleagues at the University of Iowa in Iowa City identified 16 proteins potentially derived from a photosynthetic algal ancestor. All 16 were also present in another ciliate, *Paramecium tetraurelia*.

The researchers argue that it is more likely that the ciliates had a photosynthetic ancestor but later lost the ability to photosynthesize than that they absorbed the genes by horizontal gene transfer.



A. BELL/VISUALS UNLIMITED/GETTY

MOLECULAR NEUROSCIENCE

Pores for thought

Nature Neurosci. doi:10.1038/nn.2151 (2008)

Ion channels activated by ATP — the main source of energy for cells — are abundant throughout the body, and are involved in fundamental processes such as pain and inflammation. But little is known about their molecular structure.

Shai Silberberg and her colleagues at the National Institute of Neurological Disorders and Stroke in Bethesda, Maryland, now report some unusual properties of the channels. They found that the bulk of the pore of the ATP-activated channel — and the gate controlling ion flux through the pore — comprises just three transmembrane protein stretches, one from each of the channel's identical protein subunits.

The pores of families of ion channels activated by other neurotransmitters such as acetylcholine are generally much more complex.

JOURNAL CLUB

Oliver Rando
University of Massachusetts,
Worcester

A biologist despairs over the difficulty of demonstrating heritability of chromatin states.

Chromatin, the packaged bundles of protein and DNA that make up eukaryotic genomes, is widely believed to be a carrier of ‘epigenetic’ inheritance — that is, heritable information not encoded by DNA. In multicellular organisms, the chromatin of mother and daughter cells is generally of similar shape, exposing similar regions of DNA for expression. And chromatin regulators often seem to be required for epigenetic states to be inherited. But there is a problem. It is possible that some other information carrier is inherited, and then directs chromatin regulators to re-establish a functional state.

One purported example of chromatin inheritance comes from yeast, which seem to ‘remember’ prior growth conditions. Galactose-naïve yeast induce genes for Gal enzymes slowly; those whose recent ancestors experienced galactose induce them much faster. Because this ‘memory’ requires certain chromatin regulators, it has been suggested that it provided evidence for a heritable chromatin state.

Zacharioudakis *et al.* investigate this idea using heterokaryons, fused pairs of yeast cells that have mixed cytoplasmic contents but maintain separate nuclei. By inducing memory in one yeast and seeing speedy *GAL1* expression in the other, they show that memory of galactose is transferable through cytoplasm (I. Zacharioudakis *et al.* *Curr. Biol.* 17, 2041–2046; 2007). Thus, the chromatin state around the *GAL1* gene cannot be the heritable factor, and the authors further identify the probable inheritance factor as a soluble enzyme.

These results demonstrate the difficulty of proving that any example of epigenetic inheritance is due to inheritance of chromatin state *per se*. One wonders whether any chromatin state will ever be proved to be heritable, given the difficulty of proving the absence of another information carrier.

Discuss this paper at <http://blogs.nature.com/nature/journalclub>

NEWS



B. MUHAMMAD/REUTERS

Leaders still vague on emissions targets

TOYAKO, JAPAN

World leaders met this week in Toyako on the Japanese island of Hokkaido to discuss climate change — among other global problems. But progress on reducing greenhouse-gas emissions seems increasingly unlikely to emerge from the talks.

In the run-up to this year's gathering of the G8 (Group of Eight), a clique of the world's largest developed economies, the immediate problems of soaring oil prices and global food shortages looked likely to nudge climate change off the agenda. But in fact, "food and energy issues have reinforced the climate crisis", says John Kirton, head of the G8 Research Group at the University of Toronto in Canada.

"Climate change is front and centre of the talks here in Hokkaido," says Philip Clapp of the Pew Environment Group, a non-profit organization in Washington DC, adding that Japan's Prime Minister Yasuo Fukuda "is keen to get an agreement".

But a communiqué issued by G8 leaders after a working lunch on Tuesday afternoon failed to make progress on their last meeting in Heiligen-damm, Germany, in 2007, when the G8 nations, with the exception of the United States, agreed to seriously consider reducing global emissions to 50% of 1990 levels, at least, by 2050.

Instead, leaders have declared their commitment to a vision of halving emissions by 2050, but measured against 2005 greenhouse-gas levels, rather than the United Nations standard of 1990 levels.

And despite international pressure from scientists, environmentalists and United Nations climate chief Yvo de Boer for G8 leaders to set specific, clear and nearer-term targets in Hokkaido, the latest statement recognizes "aspirational" mid-term goals, without specifying dates or levels of emissions reduction.

Most experts at the summit have been critical of the statement, and say that other G8 nations have made significant concessions to keep the United States on board. This is "a very weak, watered-down document compared with last year's statement", says Clapp. But "President Bush has now endorsed a target that the others were willing to sign up to a year ago".

"It's one step forward and one step back," says Ben Wikler of activist group Avaaz.org, which is monitoring the talks in Hokkaido. "In Bali, countries that had ratified the Kyoto Protocol agreed to work towards 25–40% reductions by 2020, and here they don't even reference specific numbers."

The United States hasn't been alone in opposing stricter targets. Although Fukuda announced at the World Economic Forum, held in January in Davos, Switzerland, that Japan would set a national mid-term goal, this is likely to be a 14% cut on 2005 levels rather than the 25–45% cuts on 1990 greenhouse-gas levels that scientists have called for.

Although historically an energy-efficient nation, Japan's emissions have still risen by more than 6% since 1990, so a 2005 baseline would mean less stringent reductions overall.

And rather than having 'top-down' targets imposed by an international body such as the United Nations, Japan has strongly endorsed taking a 'bottom-up' approach to achieving reductions by slashing emissions on an industry-by-industry basis, a method that G8 leaders recognized in the statement issued on Tuesday.

But not everyone sees the agreement that is now emerging in Hokkaido as a failure. According to Kirton, the agreement stands to "put in place a fundamentally different regime to Kyoto", one where "developing countries that are now the major carbon producers no longer get a blank cheque".

As *Nature* went to press, leaders from the G8 countries were yet to put the proposal to an additional eight 'big polluting' major-economy nations before the end of the conference. Developing nations such as China have made it clear that they want specific mid-term goals from developed countries before they sign up to halving emissions by 2050.

"There's more than enough here to reach the minimum requirements of the other eight nations," says Kirton optimistically. But Clapp isn't so sure. "Developing countries are looking for developed nations to take aggressive binding targets. This document refers to aspirational goals," he says.

Olive Heffernan

For the latest from the G8 meeting, visit <http://blogs.nature.com/climatefeedback>.

Researcher battles CNRS reforms

The French are not unaccustomed to anti-establishment protest, but the latest revolt against government reform is headed by a heroine as unlikely as Joan of Arc.

Claire Lemerrier, a 31-year-old researcher at the École Normale Supérieure in Paris, seemingly appeared from nowhere to help spearhead a movement that last week forced key concessions to a major government reform of the CNRS, Europe's largest basic-research agency.

It all began in early June, when Lemerrier heard that she was to receive the top medal for promising young researchers, an annual CNRS award. The government had just announced that the CNRS was to be broken up into six quasi-autonomous national institutes, a move widely interpreted as a dismantling of the agency.

The medal "immediately made me think of the orchestra playing on as the ship sinks", says Lemerrier. She decided to rally other medallists to protest against the CNRS reforms and

to show that the protest was not limited to left-wing trade unions and researchers scared of greater competition, but also included award-winning scientists.

Success was immediate. More than 450 medallists, including the cream of French science, have endorsed her text (<http://medailles.recherche-enseignement-superieur.fr>), which protests at aspects of the CNRS reforms, and a lack of staff and budgetary resources to implement them.

"I was pleasantly surprised at how many renowned medallists signed up — people far better known than me. I'm just the last in of the medallists," Lemerrier says.

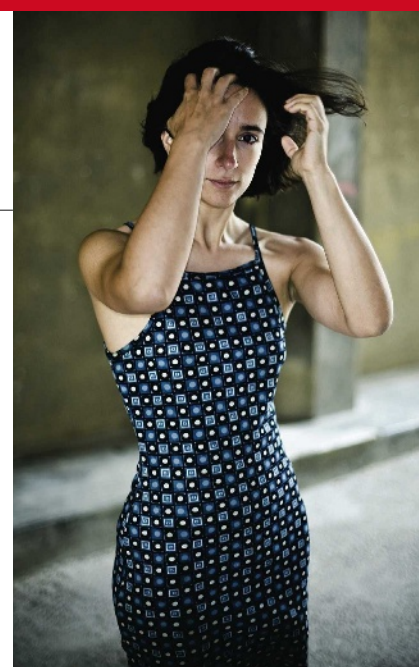
Lemerrier's initiative gave fresh and timely momentum to researchers who were already resisting the reforms and generated broad media attention. *Libération* newspaper ran a full-page profile on Lemerrier, describing her as a thorn in the ministry's side under the headline "Gunslinging chick".

The protest movement

culminated in more than 1,000 researchers invading the CNRS's Paris headquarters on 19 June and preventing the agency's board from adopting the reform. Science minister Valérie Pécresse has since made concessions that trade unions say saves the CNRS from being immediately dismantled, although they remain vigilant.

Whereas the principle of hiving off the CNRS into national institutes is retained, the revised reform would guarantee that all CNRS disciplines are represented — life sciences, the CNRS's largest discipline, as well as computing, had been relegated in the initial reform to departments. "This concession is really important, although we are still worried about how it will pan out concretely in practice," says Lemerrier.

Moreover, the CNRS, and not the government as was planned, will now appoint the heads of the institutes. This latter move had been widely perceived as an attempt by the government to take control of the CNRS. The agency will now be fully



Claire Lemerrier, 'gunslinging chick'?

F. STUCIN/MYOP

responsible for planning and managing the institutes, which will run their own laboratories. And for the first time, they will also act as research councils, giving grants to outside labs. That is something many researchers welcome, says Lemerrier.

The new reforms were adopted by the CNRS board last week, but the details will not be thrashed out until the autumn. For her part, Lemerrier says she has no ambitions to pursue her activist career further, and is keen to get back to her research.

Declan Butler

Scientists rally to Mexican researchers' plea

Scientists from around the world are calling on Mexican leaders to resolve a dispute that is tearing apart a leading South American research institute.

Two nanotechnology researchers at the Institute for Scientific and Technological Research of San Luis Potosí (IPICYT) claim they are being persecuted by the organization's administrators, after they testified in a long-running probe into the former director, José Luis Morán-López, who was ousted for nepotism in April. Morán-López denies impropriety and is appealing.

Both researchers, brothers Humberto and Mauricio Terrones Maldonado, have since lost their administrative posts. Humberto was head of the advanced-materials department, and Mauricio headed the graduate research programme. Their students claim that the current

administrators are now transferring lab facilities funded by the Terrones' grants to others.

The Terrones appealed to prominent scientists, including Nobel-prizewinning chemist Harold Kroto, one of the discoverers of 'buckyball' molecules. Last week, Kroto and 30 other scientists signed a letter in support of the pair, addressed to Mexican President Felipe Calderón Hinojosa and Mexico's national council of science and technology funding (CONACYT).

Kroto says that he fears the attacks on the Terrones will destroy one of Latin America's most promising physical-science centres and drive the two researchers abroad. "The Terrones could easily have won research positions anywhere, but they chose to go back to Mexico to help young scientists," says Kroto, who is based at Florida State University in

Tallahassee. "They have made a massive contribution to science. They now have had a kick in the teeth."

"The solidarity of the international community has been overwhelming," says Humberto. "What we want is Mexican leaders to solve the problem. Our system doesn't seem to allow for success of young scientists."

The institute's current department heads responded to the scientists' letter on 28 June. In a letter to Kroto, they decry his involvement in the affair, writing that his actions are "unbalanced, based on false premises and damage the prestige of our institute". Current IPICYT director David Rios Jara told *Nature* that the Terrones are very ambitious and "the only way they will be quiet is when they have their own centre. They don't like authority."

Rex Dalton

SPECIAL REPORT

When there's no room to grow

To maintain profits in the face of rising development costs and slow drug pipelines, big pharmaceutical firms are trying to cut back. **Heidi Ledford** examines how GlaxoSmithKline has tried to adapt.

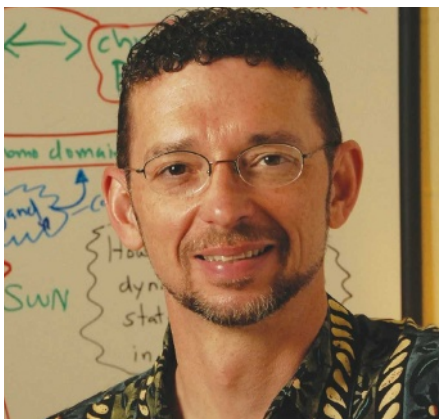
When Stephen Frye was made redundant from the world's second biggest pharmaceutical company last year it marked the end of his 20-year relationship with the firm — and perhaps the end of an era for researchers enjoying long careers in big pharma. "I grew up in that company," he says. "I had put my life and soul into it."

Frye was hired, fresh from completing his PhD in chemistry in 1987, as one of the first scientists to work at the new US headquarters of a pharmaceutical company called Glaxo. By the time he picked up his redundancy payout, he had survived two major mergers, helped discover some of the company's most prized drugs including Avodart (dutasteride), and, as vice-president of discovery medicinal chemistry, managed a group of more than 200 chemists.

Glaxo had also grown up, becoming GlaxoSmithKline (GSK), and by 2007 Frye was ready to part company. The spate of mergers was interrupting the flow of research at the firm, he says. "It was a futile cycle of constant restructuring," he claims. "A successful project takes 15 to 20 years. If you change your bets on science and people every three to four years, you will fail."

And the recurrent lay-offs — the most recent cycle of which were announced this June — drove morale to a new low. Frye decided to seek the shelter of academia, and now heads a centre for chemical biology and drug discovery at the University of North Carolina at Chapel Hill, minutes down the road from GSK's outpost in Research Triangle Park.

In many ways, the situation at GSK is emblematic of the troubled pharmaceutical



Stephen Frye spent 20 years working at GSK.

industry. The many mergers produced a drug-making behemoth with more than 100,000 employees worldwide, 17,000 of whom work in research and development (R&D). The company, like many others, faces a dilemma. To please the shareholders, GSK must continue to grow profits. But how can the company do this when it is already so large, in a regulatory environment that is reacting to a chain of drug-safety scandals, and in an industry with a high failure rate and a product cycle that lasts well beyond the attention span of the average shareholder? Some of GSK's biggest blockbuster are about to go off-patent, adding the threat of competition from generics.

Frye is quick to acknowledge that GSK is coping with these challenges to the best of its ability. But "they're struggling", he says. "Everybody's struggling. It's a tough business."

For GSK, 2007 was a tough year. In June,

Steven Nissen, a cardiovascular expert at the Cleveland Clinic in Ohio, released a meta-analysis of clinical-trial data that suggested that one of the company's biggest money-makers, the diabetes drug Avandia (rosiglitazone), could be increasing the risk of heart attack (S. E. Nissen & K. Wolski *N. Engl. J. Med.* 356, 2457–2471; 2007). Although GSK has disputed the validity of the results, the effect on sales was clear. GSK estimates that 'the Nissen paper' and the ensuing media coverage cost the company US\$2 billion.

"This [2007] is not a pretty year for us," Chris Viehbacher, president of North American pharmaceuticals at GSK, told investors at a conference last month. Viehbacher went on to assure investors that GSK was cutting costs. "There are no sacred cows," he said. "Do we have to spend \$5 billion a year on R&D? Can we do that with less?" One day earlier, news had broken that the company had laid off 350 R&D staff.

Industry-wide cuts

The lay-offs are small compared with AstraZeneca's 2007 cuts, which reduced its R&D staff by 700. Pfizer, meanwhile, laid off 10,000 employees, mostly in sales and marketing, and announced the closure of three R&D facilities in Michigan. Blogs are buzzing with rumours that additional cuts at Pfizer are on the way. Although GSK has declined to comment on specifically where the cuts were made, employees are reporting that some programmes have been hit particularly hard. More than half of the metabolic disease research group in Research Triangle Park has reportedly been let go, as have an estimated 70% of the

W. OWENS, UNC

The biotech model

It is no secret that large pharmaceutical companies have developed a voracious appetite for biotechnology firms, and GlaxoSmithKline (GSK) is no different. "I spend personally about 20% of my time shopping," Chris Viehbacher, president of North American pharmaceuticals for GSK told shareholders. The entrepreneurial spirit and nimbleness of young biotechnology companies has become the envy of the established

giants, which must cope with the inflexibility and communication woes of a large bureaucracy.

"They look longingly at these small biotech and say 'How do they do it?'" says Kenneth Kaitin, director of the Tufts University Center for the Study of Drug Development in Boston, Massachusetts. Kaitin's group has compared the cost of drug development in pharmaceutical companies and biotechnology companies and found little

difference (J. A. DiMasi & H. G. Grabowski *Manage. Decis. Econ.* 28, 469–479; 2007). "There are still people within pharma who view biotech as faster, cheaper, and more successful," says Kaitin. "But there's no evidence to support that whatsoever."

Companies often regard the 'Genentech model' as the ideal of how an arrangement between biotechnology and pharmaceutical companies should work. In 1990, Swiss pharmaceutical company

Roche bought Genentech, the flagship of the biotechnology industry, a company based in South San Francisco, California, and then essentially left the firm to function autonomously. Genentech flourished and Roche benefited, but most acquisitions have not been as rewarding. And the resignation by chief executive David Mott of MedImmune in Gaithersburg, Maryland, which was recently acquired by AstraZeneca, does not bode well. **H.L.**

**HAVE YOUR SAY**

Comment on any of our news stories, online.

www.nature.com/news

Suggestions that Avandia (inset) could increase the risk of heart disease did not make 2007 easy for GlaxoSmithKline.



cardiovascular-disease team in Pennsylvania. The GSK lay-offs are part of a three-year 'operational excellence' plan, and the company aims to institute additional cost-cutting measures in R&D over the next two years.

Some speculate that recent acquisitions may have influenced the decision (see 'The biotech model'). In April, GSK announced that it had paid \$720 million to acquire Sirtris Pharmaceuticals, a biotech company based in Cambridge, Massachusetts, that, according to Viehbacher, appealed to GSK because its drug candidates target metabolic disorders such as diabetes. In November of last year, GSK also purchased Reliant Pharmaceuticals of Liberty Corner, New Jersey, for \$1.65 billion. Reliant develops cardiovascular drugs.

When not out shopping for biotech companies, many pharmaceutical executives have been restructuring internal research programmes in an attempt to achieve the biotech firms' secret formula. In 2000, GSK research chairman Tadataka Yamada reorganized research teams into several centres of excellence for drug discovery. The teams were hived together according to the disease they were tackling, and no team was to exceed 300 people. "When you get more than that, you can't remember everybody's name," says Yamada, who now works for the Bill & Melinda Gates Foundation. In effect, the model attempts to cope with GSK's swelling size by breaking the research teams into

more manageable, biotech-like bits.

Some suggest that the recent spate of lay-offs at GSK is an indictment of the 'centres of excellence' model. Most, however, agree that it is too early to tell whether the model is a success. Robert Huckman of Harvard Business School, who has studied the model, says that several years after it was implemented GSK saw an upturn in the number of early drug candidates from the centres.

Worst of both worlds

Yamada says the 33 molecules that have emerged from the programme and already reached phase III trials are signs of success. "The model seems to be producing exactly what it was meant to do," he says. Others argue that although centre leaders held some autonomy over their research programmes, the bureaucracy remained thick. There are rumours that GSK has proposed giving its centres a three-year budget that will be renewed contingent on performance of the team, a situation much like that faced by young biotech firms. Frye worries that this creates an unstable research environment. "You get all the bureaucracy of GSK and all the risk of biotech," he says. "Three years from now, they could throw out your idea and decide to do something else."

Several other companies, including Pfizer and Abbott, have also grouped their researchers into teams organized by therapeutic class. Not everyone, however, believes that the solution to the big pharmaceutical companies'

problems is to become more like biotech companies. Two former GSK centre leaders, Frank Walsh and Robert Ruffalo, who left to join Wyeth Pharmaceuticals in Madison, New Jersey, are re-centralizing the R&D programmes there.

And researchers who have worked through several restructurings are sometimes left wondering what was so wrong with the original model. "My friends are burned out," says Paul Domanico, a former vice-president at GSK who now runs his own consulting firm. "They've been doing this kind of dance for a decade now. They truly believe if they had been allowed to focus and continue doing what they did, they would have been successful."

"You talk to any R&D scientist at a big company and you feel like you're talking to someone in Dante's Inferno."

Regardless of which model is taken, it is clear that the future will hold more lay-offs. In addition to the growing role of biotechnology acquisitions in pharmaceutical R&D, the industry is increasingly relying on contract research organizations to provide many functions that were once performed in-house. "There's a massive shift of the R&D into the smaller companies and there is going to be a massive reduction of internal R&D in big companies," says Daniel Webster of Webster Consulting Group in New York, who predicts that the waves of pharmaceutical lay-offs will continue for the next five years. The damage to morale is already palpable. "You talk to any R&D scientist at a big company and you feel like you're talking to someone in Dante's Inferno," he says. "I've never seen the morale in pharma R&D lower."

GLAXOSMITHKLINE

J. SULLIVAN/GETTY IMAGES

Q&A

A Spanish revival

Spain's Ministry of Science and Innovation was re-established in April, four years after its dissolution in 2004. **Cristina Garmendia**, a former molecular biologist and chief executive of the Genetrix group of biotech companies, heads the new ministry.

Why is the ministry putting so much emphasis on innovation?

We want to strengthen the chain of knowledge from its generation in universities and public research bodies, to the stage where it translates into a benefit to society in the form of products and services. The ministry has taken on responsibility for the main areas involved in this chain, with the goal of promoting the creation of technology-based companies. Biomedical research centres, which were managed by the Ministry of Health, are also now under the Ministry of Science.

The decision to move universities from the education ministry has been controversial. Some fear that they will be pressured by business interests.

This is not true. Of course, we have to preserve blue-sky research at universities, but we must also favour a system that makes them sustainable and improves their financing. This is not at all at odds with maximizing knowledge transfer to society. To include universities in the Ministry of Science makes sense because they have a key role in the Spanish research and development (R&D) system.

How will the long-awaited new law make Spanish science more flexible and internationally competitive?

The legislative reform will result in a new Law of Science and Technology that will give more autonomy to public research bodies and reduce bureaucracy. Changes will also be made to other laws that hinder the flexibility and internationalization of Spanish R&D.

What sort of changes?

For example, it will be simpler to apply for scientific grants. And there will be changes to immigration laws so that foreign researchers can work here more easily. Other modifications will help researchers move between the public and private sectors. We are going to introduce the new law, which has cross-party support, to the House of Representatives in the first semester of 2009.



C. CEBOLLA/ALFAQUI/NEWS.COM

One of the difficulties facing Spanish science is low investment from the private sector. How will the ministry tackle that?

At the moment, investment in R&D coming from the private sector is only about 55%, well below the European average. We hope to raise this to 66%. To do this, we must convince Spanish-based businesses, many of which are investing in R&D projects abroad, to change their practices and stay at home. We are talking with the Confederation of Spanish Industry, the Spanish chambers of commerce and employers all over Spain. To judge from the number of requests for meetings with the ministry, the private sector is very interested in the new initiatives.

How will this be encouraged?

We are offering incentives to make investment in R&D more attractive. And we are pledging full commitment in five sectors in which we wish Spain to become an international leader: health, biotechnology, nanotechnology and nanoscience, renewable energies and information technology. ■

Interview by Cristina Jimenez

See *Naturejobs*, page 248.

ON THE RECORD

“The only thing different about me is that I can’t breast-feed my baby.”

Thomas Beatie’s comment after he gave birth to a healthy daughter. Beatie, from Bend, Oregon, had a partial sex change and is legally a man. Nurses at the hospital reported that it was a ‘natural’ birth.

SCORECARD



Mercury rising
NASA’s Messenger mission has resolved a mystery about the planet closest to the Sun: its flat plains were formed by volcanic activity.



Mercury falling
Messenger also found that the radius of the 4-billion-year-old planet has shrunk by several kilometres as Mercury’s core has cooled.

NUMBER CRUNCH

10,000 people have been called in to help rid Chinese waters of an algal bloom threatening the Olympic sailing site.

285,000 tonnes of the blue-green algae have been removed from the water so far.

430 kilometres is the distance from Beijing of a potentially devastating swarm of locusts.

1 month remains until the Olympic Games begin.

ZOO NEWS

Spot the difference

A system has been developed to identify and monitor individual penguins by analysing the tiny spots on their white chests. The researchers aim to use the system on other animals too; it should work on leopards because they can’t change their spots either.

Sources: People, Science, Xinhua News, spotthepenguin.com



P.J. BARHAM

Shock tactics point to risk after quake

An analysis of the seismic stress changes in China's Sichuan basin has been published just 8 weeks after a magnitude 7.9 earthquake there killed more than 60,000 people and left millions homeless. The rapid calculation and publication¹ brings accurate forecasting of aftershocks a step closer, experts say.

Tom Parsons of the US Geological Survey in Menlo Park, California, and his colleagues describe several large faults in the basin that faced increases in stress as a result of the earthquake's main shock on 12 May. These fault sections are more likely than others to produce large aftershocks in the near future.

It is still not possible to accurately predict whether and when massive after- quakes might occur, such as the 6.0 after-shock northeast of the original epicentre that killed eight people and destroyed thousands of buildings. But experts say the study shows that it is now feasible to rapidly calculate the propagation of stress in the crust in the immediate aftermath of a major earthquake, even in areas as geologically complex as the Sichuan basin.



China devastation: measuring stress build-up can help predict aftershocks.

Such data could be used to produce large-scale maps of high-risk zones, significantly improving efforts to protect civilians in disaster regions. "We can't predict earthquakes — many scientists think we may never be able to — but we can do a lot better than nothing," says John McCloskey, a geophysicist at the University of Ulster in Colerain, UK.

Current global earthquake hazard maps are

too vague to help local emergency planners mitigate risks effectively. So a main goal of earthquake physics is to identify zones of particularly high seismic risk.

McCloskey and his colleagues analysed seismic stress perturbations in the surrounding crust after the devastating 2004 Sumatra earthquake and tsunami². Using essentially the same methods that Parsons and his team applied in estimating the stress redistribution after the Sichuan rupture, McCloskey's team successfully forecast the large quake in Sumatra in March 2005 that was triggered by the earlier rupture.

The geology of the Sichuan region is poorly understood and much more complex than the Sumatran zone, making it harder to forecast aftershocks. The seismic history of the region is also unclear. Large earthquakes are much less frequent there — every 2,000 to 10,000 years, according to an analysis published this month³ — than in Sumatra.

"It is easy to retrospectively explain why an earthquake has happened at a particular place,"

N. ELIAS/REUTERS

Super-sensitive tool key to dark-matter claim

It's one of the most controversial experiments in physics, but an Italian group's claim to have seen dark matter may be vindicated after all. A spate of theoretical papers can explain why the Italians see a signal where others do not.

Dark matter interacts rarely — if at all — with everyday molecules and atoms. Although it is thought to make up some 82% of the matter in the cosmos, scientists have so far only seen dark matter indirectly, through its gravitational pull on more conventional objects, such as galaxies.

Groups around the world are racing to spot dark matter directly, but only one claims to have actually seen it. In April, the DAMA/LIBRA (Dark Matter/Large Sodium Iodide Bulk for Rare Processes) experiment, located deep beneath Italy's Gran Sasso mountain, announced that it had evidence of dark-matter particles. The claim, the group's second in less

than a decade, was criticized for being incompatible with rival work (see *Nature* 452, 918; 2008).

"It looked at the time like they were completely inconsistent," says Kathryn Zurek, a theorist at the University of Wisconsin at Madison. But over the past three months, Zurek and other theorists have begun to find reasons why DAMA might be seeing dark matter that their rivals cannot. The papers¹⁻³ have been trickling onto the popular arXiv preprint server.

At the core of most of the papers is a previously unknown effect that DAMA claims to have seen. Known as 'channelling', it means that DAMA is much more sensitive to lightweight particles than previously thought. If dark matter were made of these lighter particles, then DAMA would see it but its rivals would not.

Just what those particles could be is at the centre of the papers now peppering

the arXiv. For her part, Zurek thinks that they could be unusually light neutralinos¹, partners of neutrinos predicted by a popular theory known as super-symmetry.

Other papers wander farther afield to explain the discrepancy. Robert Foot, a theorist at the University of Melbourne in Australia, says that the particles would be consistent with his theory of mirror matter², in which a hidden world of mirror atoms and molecules exist alongside our own. The hypothesis is part of a broader set of theories known as 'hidden-sector' models, which postulate a range of exotic hidden particles that interact with each other but not with the visible Universe.

Hints of hidden sectors would also show up in future γ -ray experiments and at the Large Hadron Collider at CERN, the European particle-physics laboratory, says Jonathan Feng, at the University of California in Irvine. "If this is the right



THE GREAT BEYOND
More 'songs about science'
can be found in our
newsblog.
www.nature.com/news

ACE STOCK LIMITED/ALAMY

McCloskey says. "But prospectively, what Parson and his colleagues have done is the best science we have at the moment."

McCloskey is setting up a group to systematically analyse changes in seismic stress in the immediate aftermath of future large earthquakes. The goal is to forecast within 12 hours the likelihood of big aftershocks, and map the high-risk 'red zones'.

"We have the methods to say strong things about what might happen days, months or years after the main shock," he says. "Aftershock forecasts will never be perfect. But in the absence of knowing something definite we're still able to advise people on the balance of probabilities."

McCloskey hopes that a scientific journal, such as *Nature*, will provide a forum for rapid dissemination of information such as daily hazard maps after a major quake.

But Philip Campbell, editor-in-chief of *Nature*, thinks such information would be better placed elsewhere because the peer-review process can slow things down. "Rushing the publication of such studies would not be good, as they usually improve quite a bit during peer review," he says. ■

Quirin Schiermeier

1. Parsons, T., Ji, C. & Kirby, E. *Nature* Advance online publication doi:10.1038/nature07177 (2008).
2. McCloskey, J., Nalbant, S. S. & Steacy, S. *Nature* **434**, 291 (2005).
3. Burchfiel, B. C. et al. *GSA Today* **18**, 4–11 (2008).

explanation, it's going to be strongly verified in the next 6–12 months," he says.

"Theoretical and phenomenological papers are always very useful," says Rita Bernabei, a physicist at the National Institute of Nuclear Physics in Rome, who heads the DAMA experiment and stands by its findings.

But other experimentalists remain sceptical. Smaller experiments using sodium iodide have failed to see the channelling effect, says Timothy Sumner, a physicist at Imperial College London. And the theoretical solutions that require lighter-weight particles do not fit with the favourite versions of supersymmetry theory. "I suspect it will be necessary to get additional data to bring the community around," he says. ■

Geoff Brumfiel

1. Petriello, F. & Zurek, K. M. Preprint at <http://arxiv.org/abs/0806.3989> (2008).
2. Foot, R. Preprint at <http://arxiv.org/abs/0804.4518> (2008).
3. Feng, J. L., Kumar, J. & Strigari, L. E. Preprint at <http://arxiv.org/abs/0806.3746> (2008).

Boyz 2 pipettemen

In a dreary, lonely lab a young female postdoc puts down her pipette to massage her aching latexed hands. Sounds like the perfect set-up for a hot new music video. Well at least it does to Tyler Kay, creative director at Compare Networks Production Group (CNP) in San Francisco, California.

A recent release from CNPG features a group of five winsome young men singing the praises of a new automated pipetting system called epMotion, made by international biotech company Eppendorf. As the lab heroine is whisked to a beach under the Golden Gate Bridge, the band members gyrate around her and her glasses are shed along with her inhibitions, just before the chorus. "Girl you need epMotion" (whispered: "yeah girl it's time to automate.")

Stefanie Noehren, online project manager at Eppendorf in Hamburg, says the company was looking for a 'viral marketing' campaign that would spread the word about the epMotion product rapidly through the Internet. And CNPG, the video production company of online biotech marketers Biocompare, was the obvious choice. In January, it created a minor internet sensation with *The PCR Song* by the mock group, Scientists for Better PCR — PCR (the polymerase chain reaction) is a lab technique used to amplify DNA. This advertisement from Bio-Rad Laboratories, based in Hercules, California, was styled after megastar group recordings of the 1980s like Band Aid's chart topping *Do They Know It's Christmas*.

In *The PCR Song*, crooners mawkishly sing their way through lyrics like: "PCR, when you need to find out who the daddy is (who's your daddy?)" In the finale, one singer lovingly smooches a thermal cycler.

"That thing took a life of its own," says Kay. It attracted more than 700,000 webpage views and spawned several homage videos from fans, singing or lip-synching the words. At Eppendorf, says Noehren, marketers decided in March to try out the music video format, this time aping one of

the prototypical manufactured male pop groups known colloquially as boy-bands — standard-bearers include 'N Sync, Take That, 98° and the Backstreet Boys.

Kay, a self-taught film-maker with Biocompare since near its inception, wrote the song with a list of product features and intense background research. "I had to listen to a whole lot of boy-band songs," says Kay. "I started to gain an appreciation for it. Those guys really know how to crank out the hits on a few chords." The result is *It's Called epMotion*, a saccharine-sweet parody of songs like 98°'s *Because of You*. Next they



Could these five men persuade you to buy their automated pipette?

COMPARENETWORKS/EPPENDORF

needed a band to sing it.

In the true tradition of boy-band manufacture, the members represent a variety of races and styles: the tough-looking Asian, the Latin lover, the bad-boy surfer, the African American with a winning smile and an odd-looking skinny one with a surprisingly deep voice. The result is a slick, if inexpensively produced video. The cost to Eppendorf for the video was just US\$50,000, says Noehren.

But does it sell the product? Amy Wagers at Harvard's Joslin Diabetes Center in Boston, Massachusetts, says "It didn't tell me much about the product, if that's what they were going for, but it's working in that now I've gotten two of these videos via e-mail." YouTube shows nearly 22,500 views and counting. More may have found the video through other sites, and Noehren says a comparable number has clicked through to the website since the video was launched in early June. "We are quite satisfied," she says. ■

Brendan Maher

Australian panel lays plans for emissions trading

An independent review of climate change in Australia has laid out suggestions for how the country might construct an emissions-trading scheme.

On 4 July, economist Ross Garnaut released a draft of the report commissioned by the Labor government of Prime Minister Kevin Rudd. Although the final report is not due until September, the draft outlines Garnaut's suggestions for how emissions trading could begin in 2010.

The report urges the government to include as many sectors as possible, including transportation, and to auction off carbon permits. Of the revenue generated, it recommends that 50% should go to households to offset resulting rises in energy prices, 30% to business, and 20% to research and commercialization of new technologies.

A government green paper outlining trading-scheme plans ahead of legislation is due this month, and will "take into account Garnaut's views", says climate-change minister Penny Wong.

UK physicists discover full scope of budget cuts

Britain will try to sell 50% of its observing time on the Gemini telescopes to help make up for an £80-million (US\$157-million) budget shortfall, the Science and Technology Facilities Council (STFC) revealed last week. Funding for astronomy, particle physics and space science are all to be cut.

The council says it will be "reviewing our future involvement in Gemini". It also slashed the UK contribution to ExoMars, a European-led mission to the red planet, by 25%. Britain had been contributing about £80 million, one-tenth of the mission's overall budget.

The STFC claimed the funding package would secure the future of Britain's Jodrell Bank radio astronomy facility, which had been in some doubt.

The council scaled down a proposed 25% reduction in UK funding for the LHCb

Iranians face nuclear-study ban in the Netherlands

The Dutch government has enacted legislation blocking Iranian nationals' access to courses and facilities related to rocketry and nuclear technology.

Under the law, Iranians living in the Netherlands, including those holding dual citizenship, will need to seek a government waiver to take courses at master's level and above in nine different subjects, including gas turbines for uranium enrichment and reactor technology. They will also be banned from the Borssele nuclear power plant (pictured) and four other nuclear facilities, says Rob Dekker, a spokesman for the Dutch Ministry of Foreign Affairs. The move comes in response to United Nations and European Union resolutions calling for restrictions on Iran's access to nuclear and weapons-related technology.



P. HILZ/HOLLANDSE HOOGTE

experiment at the Large Hadron Collider at CERN, Europe's particle-physics laboratory near Geneva. Amid concerns that such cuts would do "unacceptable damage" to the project, funding will be reduced instead by 5% in the project's first year and 10% in subsequent years.

India takes pole position to study climate change

A year after sending its first fully fledged expedition to the Arctic (see *Nature* 448, 642–643; 2007), India has established a research station in Svalbard, about 1,200 kilometres from the North Pole.

Officials signed an agreement with the Norwegian Polar Institute on 1 July to establish the base, called Himadri. The base will be managed by the National Centre for Antarctic and Ocean Research in Goa, which already maintains two research stations in Antarctica.

Research at the Arctic base will take place all year-round and will focus on climate change, officials say.

Training for human studies may become mandatory

The US government may soon require, rather than just recommend, that federally funded investigators, and the ethics-board members who approve their research, receive training and education in how to protect participants in human studies.

The Office for Human Research Protections (OHRP) announced last week that it is soliciting comments on whether such training in research ethics should be mandatory — and asking for estimates of what this would cost grantee institutions.

"Over the past several years, OHRP has identified serious, systemic noncompliance with the requirements ... for the protection of human subjects at a significant number of major institutions," OHRP regulators wrote in its *Federal Register* announcement on 1 July.

Interested parties now have until 29 September to submit their comments to humansubjectstraining@hhs.gov.

Texas educator sues over job loss and creationism

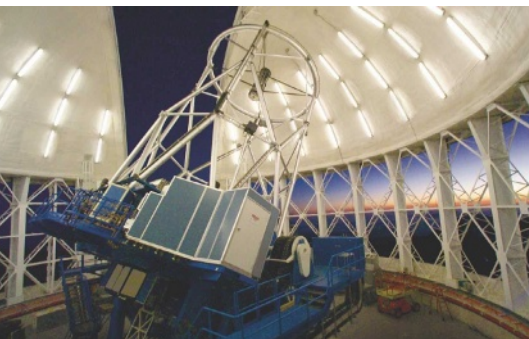
A former Texas official is suing the state's education agency, saying that its policies passively endorse creationism.

In a complaint filed with a district court on 1 July, Christina Comer, a former director of state science education, alleged that officials tacitly condone the teaching of creationism through a policy of neutrality. Comer oversaw Texas's science curriculum until last November, when she was forced to resign for circulating a notice of a talk entitled "Inside Creationism's Trojan Horse". In her termination notice, Comer was told that the education agency endeavoured to "remain neutral" on the issue of creationism.

Comer's complaint argues that board neutrality violates the separation of church and state. She is also seeking reinstatement to her former position.

NPG ARCHIVING FOR AUTHORS

Nature Publishing Group announced this week that it will provide a free service to help authors fulfil funder and institutional mandates for the archiving of primary research papers. NPG has encouraged self-archiving since 2005. The new arrangements will provide uploading for NPG authors, starting later this year. For details and to provide feedback, go to www.nature.com/press_releases/archive.html.



UK physicists face cuts on the Gemini telescopes.



THE LOST WORLD

Armed with a map depicting a 10,000-year-old landscape submerged beneath the North Sea and fresh evidence from nearby sites, archaeologists are realizing that early humans were more territorial than was previously thought. **Laura Spinney** reports.

Pilgrim Lockwood, the skipper of a British fishing trawler named *Colinda*, wasn't quite sure what to make of the thing his nets had scraped up from the bottom of the North Sea. Just over 21 centimetres long, it was made of antler with a set of barbs running along one side. Back on land, Lockwood gave the artefact to the ship's owner, and it eventually made its way to a museum in Norwich, UK. It turned out to be a prehistoric harpoon point dating to the Mesolithic period, between about 4,000 and 10,000 years ago.

That was 1931, and archaeologists studying the artefact, which became known as the Colinda point, began to realize that hunter-gatherers would once have roamed across a vast plain that connected Britain to the rest of Europe. But they had no idea what the plain looked like or what life would have been like for the harpoon's makers. Now researchers have drawn the first map of that lost world, sketching out a 10,000-year-old landscape filled with marshes, rivers and lakes. It turns out that the region they call Doggerland may have been a sort of paradise for Mesolithic people.

Because the archaeological evidence from the period is thin, Mesolithic people have in the past been depicted by researchers as restless nomads and Doggerland as a land bridge

through which they passed without leaving a trace. The new map suggests that, on the contrary, Doggerland would have been an ideal environment for them to linger in — until sea levels, rising since the end of the last ice age, finally inundated it, turning Britain into an island about 8,000 years ago. Along with other new discoveries in Britain and continental Europe, the research is helping to fill in crucial gaps in the current knowledge about Mesolithic life. "Doggerland is key to understanding the

reconstruct Doggerland in three dimensions.

In a pilot project beginning in 2002, the researchers reconstructed 6,000 square metres of the ancient landscape — slightly larger than a football field. There, about 10 metres beneath the modern seabed, they discovered the course of a major ancient river, almost as big as today's Rhine. They named it the Shotton River, after Birmingham geologist Fred Shotton who, among other things, was dropped behind enemy lines to map the geology of the Normandy beaches before the D-Day landings. Now confident that the reconstruction would

work, the researchers expanded the project. The result is a 23,000-square-kilometre map of a part of Doggerland — an area the size of Wales — that they hope eventually to extend northward as well as eastward, towards the Netherlands¹.

Archaeologists are excited about the map for several reasons. First, with an idea of how the terrain undulated, they can work out how, and how quickly, it was submerged. It is thought that the sea level rose no faster than about one or two metres per century, and that the land would have disappeared in a series of punctuated inundations. "It was perfectly noticeable in a generation," says marine archaeologist Nic Flemming, a research fellow at the National Oceanography Centre of



A sketch of the Mesolithic harpoon point found in the North Sea by the *Colinda* in 1931.

Mesolithic in northern Europe," says Vince Gaffney, a landscape archaeologist at the University of Birmingham, UK.

Along with his colleagues Simon Fitch and the late Ken Thomson, Gaffney established the mapping project to outline the terrain of Doggerland, named after the sandbank and shipping hazard of the Dogger Bank (see 'Mesolithic sites around the North Sea'). They managed to borrow seismic survey data, which outline sediment layers below the seabed, from the Norwegian oil company Petroleum Geo-Services. The researchers then put their powerful computers to work to



A Mesolithic-age artefact, made from the antler of a red deer, trawled from the bottom of the North Sea.

understand their meaning,” says Hans Peeters of the National Service for Archaeology, Cultural Landscape and Built Heritage in Amersfoort, the Netherlands.

Jan Glimmerveen, a Dutch amateur archaeologist, has over the past decade collected around 100 Mesolithic artefacts, which he gets from fishermen trawling the southern North Sea. The material includes an adze- or axe-like tool made from an antler, with part of the wooden shaft preserved, and a tool made from the bone of an aurochs, a large type of extinct cattle. Some of the artefacts have been radiocarbon dated to between 10,000 and 8,100 years ago, and all come from a small area just off the southern edge of the Birmingham map of Doggerland. The Dutch call it ‘De Stekels’ (‘The Spines’) because there are steep dunes that were probably once close to a river. Although the artefacts were lying loose on the seabed, Glimmerveen is convinced there was a Mesolithic settlement on or close to those dunes, and Peeters agrees. “You can look at it in a similar way to ploughed fields,” he says. “Objects may have been

displaced, but not over very large areas.”

Dutch researchers have a unique opportunity to retrieve Mesolithic material that was once underwater but now is part of reclaimed land. Peeters, for instance, works in a region of the central Netherlands, the Flevoland polders, that was reclaimed from an inland sea in the mid-twentieth century. In the early Mesolithic period, this area would have been tidal wetlands. Later it became a peat marsh, and the ancient bog has yielded a wealth of objects, including pottery and flint tools, that he thinks were deposited as a ritualized reference to lost ancestral lands³.

Back on the British side, archaeologists are extracting Mesolithic information from a submerged site known as Bouldnor Cliff in the Solent, the stretch of water separating the Isle of Wight from mainland Britain. As sea levels rose, swelling rivers deposited sediments over a Mesolithic valley. When the Solent began to form around 5,000 years ago, it eroded first the sediments and then the original valley floor. That erosion is continuing today, and near the Isle of Wight’s shore, it is uncovering signs of Mesolithic human activity. A few scattered flints emerged first, followed by the remains of what could be a wooden dwelling and then, last summer, tools, wood chippings and part of a log boat.

“In general the preservation is immaculate,” says Garry Momber, director of the

“Doggerland is key to understanding the Mesolithic in northern Europe.”

— Vince Gaffney

MESOLITHIC SITES AROUND THE NORTH SEA



University of Southampton, UK. “But nobody had to run for the hills.”

Second, researchers can now start to predict how Mesolithic people might have used the terrain. But it won’t be easy. Working with divers and remotely operated vehicles is complex and expensive, and the new map isn’t detailed enough for underwater excavation purposes: the smallest detectable feature on it is about 10 metres high and 25 metres wide.

To get a more detailed picture of what Doggerland might have looked like, computer specialist Eugene Ch’ng of the University of Wolverhampton, UK, is building a virtual-reality simulation. Starting with a 27-kilometre stretch of the Shotton River, he has recreated a Stone Age settlement by the water’s edge, at the confluence of two rivers, complete with thatched huts and racks for drying fish and tanning hides².

The virtual vegetation is faithful to the palaeobotanical record, right down to the stinging nettles. There are fish in the rivers, birds in the air and boar in the woods — although for now these are just avatars rather than accurate biological models. The only thing that is missing is the people, but Ch’ng will add them soon. He is constantly feeding new data into the simulation, and his ultimate goal is to turn the virtual reality into augmented reality, in which archaeologists need only don a headset to enter the Mesolithic landscape.

Found artefacts find meaning

The most immediate way in which the map will be useful, however, is in giving context to marine archaeological finds. For more than a century, fishing boats — particularly Dutch beam trawlers, whose nets scrape the seabed — have been scooping prehistoric material out of the North Sea. Most of it dates from the Palaeolithic, the vast era that ended around 10,000 years ago, and includes the bones of woolly mammoths and reindeer from the last ice age. But there is also some more recent, Mesolithic material. Until now, archaeologists haven’t been sure how to interpret these scattered remains. But with the Doggerland map, “we’ll be able to position the archaeological finds within that landscape to

Death in the Mesolithic

Mesolithic burials in northern Europe could be elaborate. People were buried lying, sitting cross-legged or on their bellies. They were buried with goods that included dogs, red deer antlers, food and amber beads. At Vedbæk Bøgebakken, north of Copenhagen, a newborn baby was buried on a swan's wing, next to a woman who is presumed to have died in childbirth.

But quite possibly, Mesolithic people in Britain didn't practise the same traditions. The only known Mesolithic burial site in Britain is Aveline's Hole, a cave in Somerset where the remains

of around 50 individuals dating from approximately 10,000 years ago were discovered in 1797. It's possible that archaeologists have been looking in the wrong places for other graves, but another explanation is that British burials were totally different from those on mainland Europe.

If so, then the submerged landscape of Doggerland could throw light on how burial practices diverged as Britain became isolated by rising seas. In Denmark, burial rites seem to have been concentrated on the coast, and there is evidence that the dead may have been placed

in canoes and sunk at sea. "I suspect that will be commonplace in Doggerland," says Kristian Pedersen, a historian at Newcastle University, UK.

Mesolithic people seem to have respected body integrity more than their successors in the Neolithic, who commonly re-entered graves to remove bones, which they then stacked in ossuaries. During the Mesolithic, "the bodies are buried fresh, they're buried intact and without a lot of disturbance", says archaeologist Liv Nilsson Stutz of Lund University in Sweden. However, there are also augurs

of what was to come. "There are examples in which people have opened the burials up and retrieved bones," she says.

Death in northern Europe was often violent. At Vedbæk Bøgebakken, one woman has a projectile point embedded in her spine, and similar finds have been made at the Breton cemeteries of Tévéc and Hoëdic. Pedersen says this fits with the notion that Mesolithic people might have embraced the concept of territoriality. "It is suggestive of some sort of conflict," he says. "People squabbling over fishing grounds, maybe."

L.S.

Hampshire and Wight Maritime Trust for Maritime Archaeology in Southampton, who is overseeing the project. He believes Bouldnor Cliff may have been a boat-building site, which is significant because it was far from the coast and so the boats would have been used only on a local lake. "We're finding evidence of sedentism," he says. "These people would have been living and working the land, maybe to a greater extent than we understand now."

Nomads settle down

Attachment to the land, ritual practices and sedentism are usually associated with later, Neolithic people. The boundary between the Mesolithic and Neolithic periods is defined as when farming begins to be practised in an area, and it generally dates to between 4,000 and 6,000 years ago in northern Europe. The stereotype of Mesolithic people is as "surviving in this harsh wilderness", says Peeters. "It was only with the introduction of farming that this poor and risky way of life was gradually brought to an end." This view, he thinks, short-changes Mesolithic people and the imaginative ways they may have used the landscape, both in life and in burial practices (see 'Death in the Mesolithic').

Evidence supporting this more complex view of Mesolithic life comes from Tévéc and Hoëdic, two Mesolithic burial sites on the coast of Brittany, France. Here, archaeologist Rick Schulting of the University of Oxford, UK, has analysed stable isotopes — mainly carbon and nitrogen — in human bones to get an idea of what the locals ate. Tévéc and Hoëdic are only 30 kilometres apart, a trivial distance for hunter-gatherers, and yet Schulting has found consistent differences in the bone isotopes between the two sites. He thinks these reflect

differences in their diet: residents of Hoëdic, for instance, seem to have got a lot more of their protein from marine resources than those in Tévéc. "That suggests to me that these people were quite embedded into the landscape over the long term," says Schulting. "They weren't moving around on a great scale."

A similar picture is emerging from Britain. In 2000, a team led by Clive Waddington of Archaeological Research Services, in Derbyshire, UK, began excavating a Mesolithic hut at Howick in northeast England. By combining radiocarbon dating with analysis of the soil strata, they were able to determine that three huts had been built at the site, each on the ruins of the previous one. Together, the huts were inhabited over about 150 years. That occupation wasn't necessarily continuous, Waddington says; nevertheless, over three or four generations people kept returning to that place.

To him, this suggests that Mesolithic people may have been staking out their group's territory. "Not that hunter-gatherers usually have any sense of ownership," he says. "But what they do have is a very strong sense of rights of access to land." Waddington argues, in fact, that the drowning of Doggerland led directly to the development of sedentism and territoriality⁴. Although the idea is speculative, it fits with the growing body of evidence for Mesolithic life in and around Doggerland. Land would have become an increasingly precious resource as the sea rose.

All these sites, taken together, may illuminate how Doggerland's residents adapted to the changing landscape. But when and how did those changes begin? The Birmingham map offers a possible clue in the shape of a giant

basin called the Outer Silver Pit, which stretches for up to 100 kilometres through Doggerland. Fed by an inlet to the east, the pit would at one time have been a lake. But two sandbanks running almost its full length could only have been formed by fierce currents. Gaffney speculates that, as the sea rose, the peaceful lake became a fast-flowing estuary into which only the most foolhardy fisherman dared launch his canoe. So what started as an attraction for water-loving people might eventually have driven them away, triggering the migration whose long-range effects Waddington is seeing at Howick.

Exact details of this upheaval will be hard to prove, not least because most of Doggerland remains uncharted. Still, archaeologists are in little doubt that such turbulent environmental change required an equally dramatic human response. In just a few thousand years, Doggerland was transformed from a harsh tundra into a fertile paradise, and eventually into the northern European landscape that we know today. "It put human adaptability to the test," says Gaffney. ■

Laura Spinney is a freelance writer in London and Paris.

"These people were quite embedded into the landscape."
— Rick Schulting

1. Gaffney, V., Fitch, S. & Smith, D. *Europe's Lost World: the Rediscovery of Doggerland* (Council for British Archaeology, in the press).
2. Ch'ng, E., Stone, R. J. & Arvanitis, T. N. *The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage* www.opennature.org/Portals/0/Documents/ShottonRiver.pdf (2004).
3. Peeters, H. *Hoge Vaart-A27 in Context: Towards a Model of Mesolithic-Neolithic Land Use Dynamics as a Framework for Archaeological Heritage Management* (RACM, Amersfoort, 2007).
4. Waddington, C. (ed.) *Mesolithic Settlement in the North Sea Basin: A Case Study from Howick, North-East England* (Oxbow Books, Oxford, 2007).



THE BRAINS OF THE FAMILY

Does the difficulty in finding the genes responsible for mental illness reflect the complexity of the genetics or the poor definitions of psychiatric disorders? **Alison Abbott** reports.

Every family has its foibles, but this one has more than most. The first member came to researchers' attention in 1968 as part of a genetic survey of juvenile delinquents who had been admitted to Scottish detention centres. One boy carried a major 'translocation', in which a chunk of chromosome 11 had been switched with part of chromosome 1.

The translocation and the boy's bad behaviour were more than just coincidence. Years later, when Edinburgh researchers traced the family, they found that the same chromosomal abnormality spanned four generations, with remarkably varied effects. Of those who carried it, five had depression, six had schizophrenia or related disorders, three had adolescent conduct disorder and two had anxiety disorder. One had attempted suicide and died in a mental hospital¹. Several of those without the translocation had their own problems, including anxiety, minor depressive disorder and alcoholism.

The translocation chops up a gene called, with some justice, disrupted-in-schizophrenia 1 (*DISC1*)². This gene is arguably the strongest contender for involvement in any psychiatric disease, and a flurry of activity has shown that it codes for a hub protein involved in multiple biochemical pathways in the nervous system. But the association and its Scottish heritage are also a textbook example of the complexity of psychiatric genetic studies. *DISC1* has not shown up in many other screens for schizophrenia genes,

even though it has occasionally been identified in genetic studies of other illnesses, including bipolar disorder and autism.

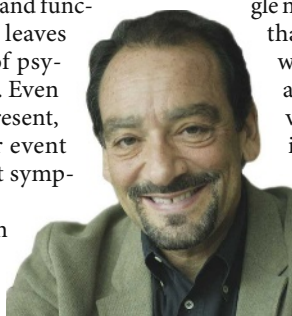
Researchers are convinced that variations in *DISC1* and many other genes can scramble the intricate and as yet inscrutable processes by which brain circuits develop and function, and that this scrambling leaves some people at greater risk of psychiatric problems than others. Even if one or more risk genes are present, external stress or some other event may decide whether and what symptoms are triggered.

Finding genes involved in psychiatric conditions is proving to be particularly intractable because it is still unclear whether the various diagnoses are actually separate diseases with distinct underlying genetics or whether, as the *DISC1* story suggests, they will dissolve under the genetic spotlight into one biological continuum. Indeed, some researchers suggest that it would be better to abandon conventional clinical definitions and focus instead on 'intermediate phenotypes', quantifiable characteristics such as brain structure, wiring and function that are midway between the risk genes involved and the psychopathology displayed.

In the past two years, researchers have pulled out a host of genes involved in other multifactorial diseases, such as diabetes and obesity, by use of genome-wide association studies. These use powerful new genomic tools to scan for variations in the DNA sequence called single nucleotide polymorphisms (SNPs) that tend to occur in individuals with a particular condition. They allow scientists to see which gene variants pop up more frequently in people who have a disorder.

Finding small genetic signals is a question of statistics: a weak association between a gene and a disease may stray into significance only when a study has hundreds or thousands of participants. But instead of helping to firm up which genes might be candidates, the largest population studies completed so far in psychiatric genetics seem to be eliminating them. A study this year led by Patrick Sullivan,

a geneticist at the University of North Carolina at Chapel Hill, involved nearly 750 patients with schizophrenia and a similar number of controls, and analysed almost half a million SNPs. But not one gene met the rigorous statistical requirements needed to show it was a risk factor³ — not even *DISC1*.



**"Geneticists know nothing about psychiatric disease."
— Daniel Weinberger**

deny that it is a crude approximation and that the symptoms used to assign categories substantially overlap. Delusions are frequently a symptom of schizophrenia, but a typical person with schizophrenia might also exhibit apathy. Patients with major depressive disorder usually exhibit apathy, but some will also be delusional. A clinical diagnosis “may be the mental end-stage of all sorts of different neurobiological disorders”, says Robin Murray, a psychiatrist from Kings College London. “It may be a little like a diagnosis of ‘renal failure’ which is the end-stage of all sorts of medical disorders from heart disease to toxic shock.”

This presents a profound problem for genetic studies because it means that some researchers may be looking for shared risk genes in a group of people who may not actually share any or, indeed, have the same disease. It is an absolutely crucial point; many researchers think that poor and inconsistent diagnoses may explain much of the past failure of genetic-association studies and will be the factor on which future success hinges. The difficulties became apparent even before the genome-association studies. The *DISC1* story — in which a strong candidate identified in one study is rarely found in another — has been repeated over and over in family studies and in case-control studies, in which groups of unrelated individuals with a disorder are compared with healthy controls.

Weinberger is convinced that the right signals will be found by his candidate-gene approach, and says that millions of the Lieber family's money will be spent to that end. Weinberger likens the huge endowment to the Manhattan Project in its ambition and scope. “We'll try to work out the risk genes, work out the faulty circuitry and work out how to develop new ways of treatment,” he says.

Weinberger's first major breakthrough involved a gene called *COMT*, which encodes a key enzyme in the metabolism of dopamine, a neurotransmitter that is important in schizophrenia and the target of many antipsychotic drugs. *COMT* had already been linked to schizophrenia in family and case-control studies. But Weinberger instead examined whether variation in the gene was linked to one of his intermediate phenotypes. In this case he used functional magnetic resonance imaging (fMRI) to examine the activity of the brain's prefrontal cortex, which is known to function abnormally in schizophrenia.

In a 2001 paper, Weinberger and his colleagues showed that people with a variation in



COMT that causes a methionine amino acid to be substituted for a valine in the protein tended to have worse working memory and lower activity in the prefrontal cortex than those without the mutation. They also found that the mutated version of the gene and the poorer working memory were found in more parents whose offspring had schizophrenia than in those who didn't⁴. That makes biological sense, says Weinberger, because the two forms of the

enzyme differ in how efficiently they break down dopamine, so children who inherit the methionine version would have upset levels of a key transmitter that might make them more susceptible to schizophrenia.

The study was a turning point in the field, says Jeremy Hall, a neuroscientist at the University of Edinburgh, UK, who was one of those convinced to adopt the approach.

But Flint argues that no one has yet shown that intermediate phenotypes have a tighter link to the genes responsible than the disease itself. “I just don't buy that brain size or whatever will work better,” he says.

Weinberger quickly followed his 2001 paper with another in which he examined variants of the serotonin transporter gene that are known to be risk factors in psychiatric disorders such as depression. Rather than looking for an association between genetic variants and a specific disorder, Weinberger looked for links with another intermediate phenotype. He used fMRI to show that those with the ‘short’ version of the gene — who have less efficient serotonin signalling — also had more activity in the amygdala when they reacted to fearful

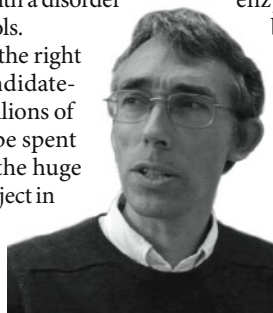
stimuli such as angry or afraid faces⁵.

It was the first time that anyone had shown that a gene could change how a region of the brain responds to an emotional stimulus — the idea being that it could predispose someone to a range of psychiatric conditions if he or she encounters the right environmental trigger. In 2003, a study done in New Zealand independent of Weinberger identified one of those potential genes. The team used data from a long-term study that tracked 1,000 children through young adulthood, and showed that those with the short variant of the serotonin transporter gene were more likely to be thrown into serious depression when faced with a stressful life crisis, such as relationship angst, than were those with the long variant⁶. It's a classic case of gene-environment interaction, and Weinberger has gone on to identify other gene variants, such as one that could make children more likely to develop schizophrenia if they are exposed to a period without oxygen during birth.

Emotional response

The main criticism of the candidate gene approach is that it is restricted to the tiny list of genes that have established links to the disease, precisely what researchers are short of. “We are just too ignorant of the underlying neurobiology to make guesses about candidate genes,” says neuroscientist Steven Hyman, provost of Harvard University in Cambridge, Massachusetts. “Candidate genes are like packing your own lunch box and then looking in the box to see what's in it.” That's why Hyman and many other researchers say that, despite their past failures, bigger and better genome-wide association studies are a more promising way forward for psychiatric genetics.

Several meta-analyses are now under way in attempts to increase the statistical power and



“The studies so far are statistically underpowered. We need bigger studies.”

— Jonathan Flint

CORRESPONDENCE

International law should govern release of GM mosquitoes

SIR — Your News story ‘Sterile mosquitoes near take-off’ (*Nature* **453**, 435; 2008) discusses the likely release of genetically engineered mosquitoes to help contain dengue fever. It demonstrates just how close we are to a radically new set of strategies for managing a whole range of diseases and wildlife using genetically modified organisms (GMOs). But after assessing the risks and benefits, nations may reach different conclusions about their use. And that’s quite a problem, considering that genetically modified bugs won’t recognize national borders.

Malaysia may successfully avoid spreading the sterile mosquitoes across the border, and even if they do cross, any transgression will be limited to a single lifespan. Strategies that rely on genetically modifying mosquitoes that can still reproduce, such as those that you mention are being engineered by Austin Burt’s team for malaria control, are more likely to have lingering effects as they spread among wild populations across national borders.

Although everyone would agree that diseases are bad, nations may have very different views about how they should be prevented. Effective control of mosquito-borne disease will not be achieved without regionally coordinated programmes, so a robust framework is needed to accommodate the differences (B. G. J. Knols *et al.* *Am. J. Trop. Med. Hyg.* **77**, 232–242; 2007).

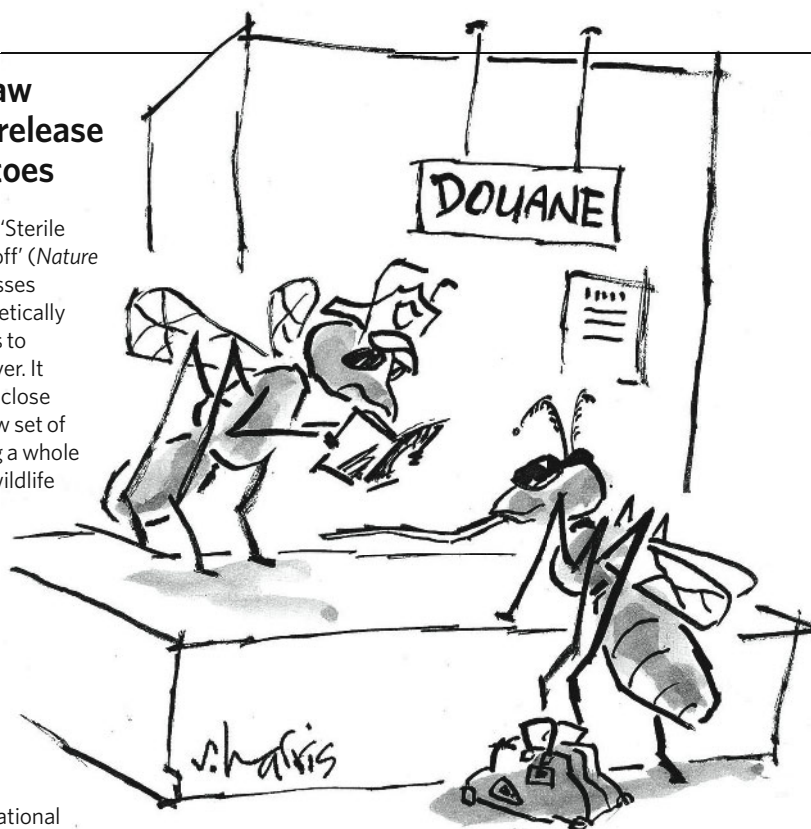
Nations have a right to decide the technological risks to which they expose themselves. The factors in decision-making here will not be only the simple ones of uncontested science — this is

politics, and appropriately so. The potential for conflict over self-dispersing GMOs demands the attention of international law.

Unfortunately, that law is deficient. The most relevant treaty is the Convention on Biological Diversity and its instrument, the Cartagena Protocol on Biosafety. The protocol includes Malaysia and most of its neighbours — Thailand, Indonesia, Cambodia and Vietnam — but not Singapore or Australia. It falls far short of the proactive and instructive approach necessary for the deployment of GMOs designed to be self-dispersing and reproducing in the wild.

The world needs to get involved. Discussions should be formalized, ideally under the biodiversity convention, and widely accessible (for example, through the convention’s Biosafety Clearing House). This trial would set the tone for future negotiations over this very new kind of biotechnology.

Elena Angulo Laboratoire Ecologie, Systématique et Evolution, UMR CNRS 8079, Université Paris Sud, 91405 Orsay Cedex, France, and Doñana



Biological Station, CSIC, Apdo. 153, 41080 Sevilla, Spain
Ben Gilna Department of Geography, University of Hull, Hull HU6 7RX, UK

Future of open access could be online and peer-reviewed

SIR — Raghavendra Gadagkar (*Nature* **453**, 450; 2008) argues that the open-access ‘pay to publish and read for free’ model leads to a disadvantage for scientists in developing countries. I disagree. Gadagkar correctly states: “page charges may be waived for authors who cannot afford to pay.” He then adds: “a model that depends on payment by authors can afford only a few such waivers.” This is not necessarily true: for example, some open-access journals provide discounts to particular institutions.

I would prefer to see what little money is available to a developing country spent on helping to publish their scientists’ papers rather than financing publishing

houses based in First World countries. At present, open-access publication may be hard for those in the developing world to afford, but in the long run it will be advantageous, offering them free access to educational and academic resources.

Most important, the future of open access probably does not lie in journal publishing models. The huge success of online literature databases such as arXiv (<http://arxiv.org>), free to publish and access, is significant. Such databases currently host mostly non-peer-reviewed preprints, and so are of little value for career building. But academic organizations throughout the world could, if they wished, build an equivalent archive of peer-reviewed papers.

I also disagree with Gadagkar’s view: “If I must choose between publishing or reading, I would choose to publish”. No one can expect to do serious science without access to the current academic literature.

Although many subscription journals are free to access online in developing countries through the HINARI, AGORA and OARE initiatives of the United Nations, the principle remains that if you cannot afford to read, you automatically cannot afford to publish. Perhaps Gadagkar will agree next time he is denied access to a fundamental paper for his research because his institution does not subscribe to it.

Massimo Sandal Department of Biochemistry G. Moruzzi, University of Bologna, Via Irnerio 48, 40126 Bologna, Italy

Genetic testing must recognize impact of bad news on recipient

SIR — Your Special Report ‘Genetic testing for everyone’ (*Nature* **453**, 570–571; 2008) discusses the contentious issue of breaking the news about test results. Our experience

“The CDZ model could guide future efforts to explain the relationship between what we see and what we do.” see page 167

of conducting genetic-testing enquiries in Colombia, which has the largest kindred in the world with familial Alzheimer's disease, indicates that individual resilience may vary greatly. Justifying relaxation of recommendations for vigilance on the basis of findings “that most people are remarkably resilient in the face of traumatic genetic test results” gives short shrift to those who are not.

When we asked a 24-year-old man whose mother harbours the highly penetrant presenilin mutation what he would do if he tested positive for the gene, he indicated that he would shoot himself in the head. Although he would probably have another 24 dementia-free years ahead of him, because the average age of disease onset in this community is 48 and variation around the mean is relatively small, the eventuality of a disease many years in the future pervaded his thinking.

Seeking predictive genetic testing can be a risky behaviour, and an individual's likely response to genetic risk is hard to foretell. Functional magnetic resonance imaging activity patterns may be able to define people who are more comfortable with risk, and genetic polymorphisms seem to contribute to risk-taking behaviour. Defining the scientific basis for how individuals handle volatile genetic information may help guide our decisions about the best setting for delivering predictive-testing news.

At what point does genetic destiny overtake the hope of beating the odds? In the Colombian families, an affected parent already sets the risk of disease at 50%, a level that in our experience creates significant anxiety but is tolerable.

Genetic testing, whether it offers a stick of dynamite or a stark warning to which we can adapt, must be backed up by reliable, accessible, up-to-date information. For example, the completion of phase III clinical trials can radically alter a bleak

message detected in the genome.

Kenneth S. Kosik Neuroscience Research Institute, University of California–Santa Barbara, Santa Barbara, California 93106-5060, USA
Francisco Lopera Grupo de Neurociencias de Antioquia, Universidad de Antioquia, Medellín, Colombia

Action needed to prevent extinctions caused by disease

SIR — Your News in Brief item ‘Cancer forces Tasmanian devil onto endangered list’ highlights the plight of this carnivorous marsupial (*Sarcophilus harrisii*), driven towards extinction by devil facial-tumour disease, which is contagious (*Nature* **453**, 441; 2008). The animal will soon also be uplisted by the 2008 IUCN Red List from its category of Least Concern to Endangered.

Emerging disease has also had a sudden impact on the western gorilla (*Gorilla gorilla*), uplisted in 2007 from Endangered to Critically Endangered, primarily as a result of mortality (more than 90% in some remote areas) induced by Ebola virus (*Nature* **449**, 127; 2007). Disease has caused a 33% decline in the gorilla population over 13–14 years and a 64% decline over 11 years in the devil population, and is set to continue. Such rapid, range-wide population crashes have also been documented in formerly widespread and common amphibian species affected by the pathogenic chytrid fungus (*Nature* **439**, 161–167; 2006).

The multimillion-dollar Save the Tasmanian Devil programme was initiated by Tasmania's Department of Primary Industries and Water in 2003. By contrast, nascent efforts to vaccinate wild gorilla populations with newly developed vaccines against Ebola virus are meeting with resistance from some members of the conservation community, who fear negative impacts during the vaccine-testing process.

The lesson to be learned from each of these diseases is that, although aggressive management actions, such as vaccination, may negatively affect a handful of individuals, they are vital if we are to save entire species.

Michael Hoffmann IUCN/SSC-CI/CABS Biodiversity Assessment Unit, c/o Conservation International, 2011 Crystal Drive, Ste 500, Arlington, Virginia 22202, USA
Clare E. Hawkins Threatened Species Section, Department of Primary Industries and Water, GPO Box 44, Hobart, and School of Zoology, University of Tasmania, Private Bag 5, Hobart, Tasmania 7001, Australia
Peter D. Walsh Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany

Cuddly animals don't persuade poor people to back conservation

SIR — In your Editorial ‘Two symbols, one solution’ (*Nature* **453**, 427; 2008) on symbols used to publicize the challenges of global warming, you caution against focusing on animals rather than people. But symbols can be powerful — so perhaps what we need instead is to identify a new set to serve the interests of environmental conservation.

The appeal of animals and idyllic forests is not universal. Criticizing this Western imagery, which seems to rate animals more highly than people, has become a rallying cry for local leaders the world over. It tarnishes conservation as a ‘new colonialism’. For example, a candidate for the governorship of East Kalimantan (Indonesian Borneo) recently declared that people should take precedence over orangutans (*Pongo pygmaeus*) — a politically pragmatic opinion that is unlikely to change soon in a country where recent figures indicated that 52.4% of the people live on less than US\$2 per day.

Here in Indonesia, iconic images show the noble fight

against poverty. Again, wild animals are largely irrelevant. Local community members on one of our orangutan conservation programmes were puzzled as to why we didn't help them first — “for we are the *orang utan*” (‘forest people’).

People in developing countries are seldom against conservation itself, although they may resent the conservation imposed on them. The hard slog of putting conservation into practice — economic planning, land-use allocation, calculation of environmental services, policies, sustainable financing and law enforcement — must be translated into stories and symbols that translate across cultures so that they can be better understood.

The dollar sign, for example, could symbolize opportunity fees that the wealthy would be prepared to pay to implement their own brand of conservation. A ballot paper could signal that conservation is subject to the same democratic checks and balances that we require in other societal choices. Such abstract, process-based ideas are less photogenic than cuddly animals, but they could potentially be key to reducing loss of tropical forests and thereby saving orangutans and other threatened species.

Erik Meijaard The Nature Conservancy Indonesia Forest Program and the Orangutan Conservation Services Program, Markoni, 76112, Balikpapan, East Kalimantan, Indonesia
Douglas Sheil Center for International Forestry Research, PO Box 6596 JKPWB, Jakarta 10065, Indonesia, and Institute of Tropical Forest Conservation, PO Box 44, Kabale, Uganda

Correction

In J. M. Swanson and N. D. Volkow's Correspondence (*Nature* **453**, 586; 2008), consumption estimates of stimulant drugs by country in 1995–2006 were wrongly attributed to the World Health Organization. These data were collected for the United Nations by the International Narcotics Control Board, and were taken from the published annual reports.

COMMENTARY



Viral strains injected into chicken eggs still form the basis of the annual flu vaccine. It is a six-month process and up to two eggs are required for each dose.

B. INDAHONO/EPA/CORBIS

The contents of the syringe

The influenza vaccine failed this winter. **Steven Salzberg** suggests that future success relies on sharing data more widely and making the virus strain selection process more transparent.

During the past 50 years, the scientific community has studied the influenza virus in great detail, and has developed an effective vaccine that is administered widely each year. The vaccine contains isolates from each of the three strains commonly circulating in humans: H3N2, H1N1 and influenza B. H3N2 has been the dominant strain of influenza A in most years, since it first emerged in 1968, and it is responsible for the most serious infections. The milder H1N1 has been co-circulating since it reappeared in Russia in 1977, and influenza B is milder yet. Because the virus mutates rapidly, the vaccine strains — especially H3N2 — need to be changed almost every year in order to remain effective. In some respects, the influenza-vaccine programme is a remarkable success: every year a new vaccine is developed and distributed, and most of the time it works.

This year, however, the vaccine was a failure: the strain of H3N2 that was used provided very little protection from infection. After a mild start dominated by H1N1 a new type of H3N2 emerged in mid-winter and quickly dominated, soaring to 71% of cases in the first 8 weeks of 2008 and overwhelming medical clinics

in many places. A study in Wisconsin found the vaccine to be only 44% effective compared with the 70–90% effectiveness expected¹, and a Harris Poll of more than 2,500 people revealed that for the first time in at least four winters, people who were vaccinated seemed no less likely to become infected². The harm was thus twofold; people fell ill and their trust in the vaccine system was undermined. This failure could have been predicted, if not prevented, through a more open system of vaccine design, a stronger culture of sharing in the influenza research community and a serious commitment to new technologies for production. The habits of the vaccine community must change for the sake of public health.

What happened?

Twice each year, the World Health Organization (WHO) organizes a meeting with the directors of the four WHO Collaborating Centres and three Reference Laboratories to make recommendations for the composition of the influenza vaccine. The experts convene in February to determine the vaccine for the Northern Hemisphere, and in September for the Southern Hemisphere. The WHO publishes

its recommendations immediately afterwards. In February 2007, the recommended strains for the 2007–08 season were: for H3N2, a 2005 strain from Wisconsin; for H1N1, a 2006 strain for the Solomon Islands; and for influenza B, a 2004 strain from Malaysia.

H1N1 dominated the 2006–07 season, and the WHO panel decided — correctly, it turned out — that the 1999 strain in that season's vaccine was no longer effective. They also looked at surveillance data on H3N2, which showed that a new strain of H3N2 was emerging. According to a WHO report, "an increasing proportion of recent viruses showed antigenic differences from the vaccine virus"³. However, there was some uncertainty about what the new H3N2 strain would be, and furthermore "the lack of egg isolates precluded the selection of a new vaccine candidate". In other words, the strains that seemed to match the most recent H3N2 cases could not be adapted to grow well in eggs. The committee therefore chose to take no action: they recommended that the vaccine continue to use the 2005 Wisconsin strain.

This strategy was likely to fail, given that the WHO already had evidence that H3N2 was changing. Had the season's influenza A

infections been caused mainly by H1N1, few would have noticed the H3N2 mismatch outside the influenza research community. But from January 2008, the profile changed, and H3N2 isolates dominated for the rest of the flu season (see graph).

The choice of the wrong strain for the vaccine had dire consequences: in addition to a rapid increase in H3N2 in early 2008 (a 17-fold increase in the first eight weeks of the year), pneumonia and influenza mortality exceeded epidemic levels — defined as 6.3% of all deaths reported to the Centers for Disease Control and Prevention (CDC) — for 19 consecutive weeks starting with the second week of 2008, according to CDC data from 122 US cities.

A closed process

The WHO met on 11–13 February this year to decide on the strains to be included in the vaccine for the 2008–09 season. As usual, the meeting was closed to all but invited participants, who this year included members of the WHO influenza surveillance network, representatives of national drug regulatory agencies, and influenza vaccine manufacturers⁴. The experts involved chose to replace the H3N2 strain with a more recent isolate, from 2007, which should be a better match to the circulating viruses next season. Neither the WHO nor the CDC publishes the evidence used to support their decision. That evidence includes hemagglutinin inhibition tests of hundreds of isolates, genome sequences of some isolates and data on the ease with which the isolates can be grown in eggs.

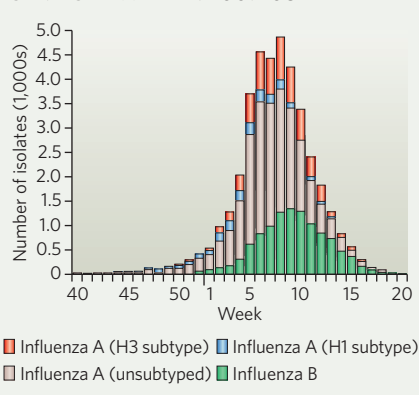
The process of choosing flu-vaccine strains needs to be much more open. Other scientists, such as those in evolutionary biology with expertise in sequence analysis, could meaningfully contribute to the selection. At present, external scientists cannot review the data that went into the decision, nor can they suggest other types of data that might improve it.

The selection of a vaccine strain has tremendous public-health consequences. Such a decision should be described in at least as much detail as a scientific publication. Although time constraints make formal peer review impractical, at a minimum all the data used to justify the selection should be immediately made public. The WHO and the CDC should encourage other scientists to re-analyse the data themselves and to critique the decision.

The leaders of the influenza community, especially the WHO and the CDC, should create policies — for sharing data and isolates — that are more open, and should insist that their own scientists follow those policies. When these leading organizations set an example, the rest of the community will follow.

Of course, preparing vaccine in cell culture could reduce some of the pressures put on that dark room of vaccine predictors. The current system, in which most of the world's vaccine supply is grown in chicken eggs, is an antiquated, inefficient method requiring six months or more to ramp up production, which in turn means that the vaccine strains must be chosen far in advance of each flu season. More crucially it sometimes prevents the use of the optimal strain, as it did in 2007. And, if the next pandemic is an avian-influenza strain such as H5N1, then it could easily sweep through the chicken farms that we rely on to produce eggs for vaccines. Cell-culture methods for influenza vaccines are under development in Europe and the United States, with strong

INFLUENZA CASES IN THE UNITED STATES — WINTER 2007–08



government support — the National Institutes of Health (NIH) awarded a contract to pharmaceutical company Sanofi Pasteur in 2005 to work on these — but the investment required to gain approval remains a significant hurdle. These efforts need to be accelerated if we are to be ready for the next pandemic.

Share sequences and isolates

Genome sequences of currently circulating isolates, both human and avian, are among the most important sources of information about influenza. Despite growing calls for the rapid release of influenza sequence data, the practices of most major influenza research groups — including the WHO and the CDC's own influenza centres, and most large, NIH-funded influenza centres in the United States — have been slow to change. Scientists at these centres hold onto biological samples and sequences for as long as possible, usually until a scientific article is published, and sometimes indefinitely.

This practice stands in stark contrast to the genome-sequencing community, which realized at least ten years ago that society reaps the greatest benefit when scientific data are shared

rapidly and openly. Rapid sharing was a major tenet of the Human Genome Project, and it has become a requisite for publicly funded research in the United States and United Kingdom. In 2004, in an effort to establish a similar model for influenza, David Lipman and I started the Influenza Genome Sequencing Project, an NIH-funded effort dedicated to sequencing as many influenza genomes as possible, with the goal of dramatically expanding our knowledge of influenza epidemiology and evolution⁵. The project has just finished its 3,000th genome, and all data have been deposited in GenBank, the publicly available NIH genetic sequence database.

The WHO and the CDC have stated publicly that they support placing sequence data in the public domain^{6,7}. Unfortunately, the WHO's own centres do not release all their influenza sequences, and when they do, they often use the Los Alamos National Laboratory influenza database. This database is, as reported on its own website, "a private database for collaborators" — access is restricted to a private group of subscribers. A closed database limits the free exchange that is so important to scientific research, and it sets the wrong example.

The database created recently by the Global Initiative on Sharing Avian Influenza organization (GISAID, www.gisaid.org) might represent a new model for influenza data sharing. GISAID allows registered users to have access to all its data with a few restrictions, designed to address legitimate concerns that some scientists and governments have expressed: users must agree to acknowledge the data providers, and they are expected not to make intellectual-property claims that might tie up the data or restrict use for others. The database is still too new to know if it will succeed in its goal of encouraging greater data sharing, but if it does, it could be an important step forward.

We can't control how the influenza virus mutates from year to year. But we can do much more to track those mutations, and by sharing what we learn, we can increase our odds of beating the flu.

Steven Salzberg is at the Center for Bioinformatics and Computational Biology, 3125 Biomolecular Sciences Building, University of Maryland, College Park, Maryland 20742.

1. www.cdc.gov/mmwr/preview/mmwrhtml/mm5715a1.htm
2. www.harrisinteractive.com/harris_poll/index.asp?PID=906
3. www.who.int/csr/disease/influenza/20078anorthreport.pdf
4. www.who.int/csr/disease/influenza/vaccine_north2008_9/en/index.html
5. Ghedin, E. et al. *Nature* **437**, 1162–1166 (2005).
6. www.who.int/csr/disease/avian_influenza/guidelines/h5n1sequences2006_08_23/en/index.html
7. www.cdc.gov/od/oc/media/pressrel/r060822.htm

See Editorial, page 137. To discuss this article online visit <http://tinyurl.com/625swp>.

SOURCE: WHO/NREVS

COMMENTARY

Ready for avian flu?

Committing to a vaccine stockpile is just the beginning. **Tadataka Yamada, Alice Dautry and Mark Walport** offer a roadmap for heading off a global avian influenza catastrophe.

Several chilling considerations highlight the seriousness of an impending pandemic of the H5N1 'avian' influenza virus. The highly contagious nature of influenza, the limited ability to restrict its transmission and the efficiency of modern international transport all conspire to reduce the time from the first infection to a potential global crisis. Recent models built on data from the 1918 flu pandemic predict that 50 million–80 million people could die¹. Perhaps not surprisingly, 95% of these deaths are likely to occur in the developing world, where higher population density, poor health status and limited access to public-health interventions prevail. Prevention through vaccination would be optimal, but vaccines against a pandemic strain might take six months to manufacture and deliver, even in developed countries. Moreover, total global capacity for flu vaccine manufacture in the first 12 months is estimated at only 500 million doses, and no global financing-vehicle exists.

There is some good news. The World Health Organization (WHO) has announced plans to stockpile H5 influenza vaccine and several manufacturers have already offered to contribute. Other manufacturers are supportive but await details before committing. Efforts have been initiated by the WHO to create a policy framework for vaccine allocation and recommendations for its use. An ethics framework, and financing, regulatory and distribution systems will also have to be developed with member states.

Several recent developments make this stockpile feasible. H5N1 vaccines with adjuvants that reduce the required dose as much as fourfold have been developed² and one has been licensed for medical use. Furthermore, the manufacturing capacity of 500 million doses is calculated on a requirement for three strains of flu virus for standard vaccinations; in crisis mode, three times as much monovalent pandemic flu vaccine could be produced. Together, these considerations could increase global vaccine production capacity to 5 billion–6 billion doses over 12 months. Moreover, adjuvant-enhanced vaccines may provide cross-protection against strains that have undergone up to seven years of genetic drift³. If this is true, appropriate planning, manufacture and stockpiling of currently effective vaccines might provide the basis for an immediate response to an H5N1 outbreak.

But this is not the time to be complacent.

Our organizations, the Bill & Melinda Gates Foundation, the Pasteur Institute and the Wellcome Trust, are now planning, with major medical-research funders and other stakeholders, several projects to enhance the research effort and reduce the risks from the threat of pandemic influenza over coming decades. In the next 18 months we will develop, maintain and disseminate a central inventory of funded research activities that are relevant to human



Vaccines: still the best defence against avian flu.

influenza to ensure that stakeholders are well-informed. We will also coordinate roadmapping exercises to identify knowledge gaps. These will assist funders and researchers in establishing research-funding priorities, with specific focus on vaccines, drug therapies and epidemiology/population science (for example, diagnostics, surveillance, transmission and modelling). The Bill & Melinda Gates Foundation and the Wellcome Trust will collaborate to fund these activities.

These two programmes will enable the development of a cohesive health-research agenda for pandemic influenza. In addition, we call on the wider community to focus on the following challenges.

Resolve intellectual-property claims

We, as a community, must be transparent and open in sharing information, reagents, viral strains and scientific know-how to conduct and advance the best science and rapidly create the most effective pandemic vaccines. We must ensure that intellectual property concerns in either the developing or the developed world do not stand in the way of effective solutions.

Coordinate stockpiling efforts

We must seize the opportunity afforded by emerging science to mobilize the global capacity to manufacture an effective 'pre-pandemic'

H5-vaccine stockpile with novel vaccines currently under evaluation. Although the capacity to manufacture pandemic flu vaccines is large, we must continue to supply standard seasonal flu vaccine to those at risk.

Consider all costs

We must make sure that all people from all nations will benefit from stockpiling, while ensuring fair financing. This implies a tiered-pricing model across countries. Careful consideration must also be given to providing adequate protection for health providers and manufacturers against costs of claims for unpredictable side-effects of vaccination.

Initiate a funding mechanism

Even at the lowest prices it will be difficult for many nations to meet the needs of their people. Public and private donors must create a financing facility to help the poorest nations pay for their share of the pandemic flu vaccine. Donors have funded financing mechanisms, such as the Geneva-based GAVI Alliance, which supports the purchase of childhood vaccines, to address other urgent health challenges.

Bolster surveillance

We must build robust mechanisms for surveillance of an outbreak of pandemic flu and for delivering prevention and treatment, particularly vaccines, as quickly and broadly as possible. There must be full integration of vaccine strategies with other approaches, and we must coordinate research strategies for dealing with zoonotic and human influenza infections.

Advances in science and society can help prevent disaster if we resolve to act collectively. ■ Tadataka Yamada is the executive director of the Bill & Melinda Gates Foundation's Global Health Program, 1551 Eastlake Avenue East, Seattle, Washington 98102, USA. Alice Dautry is director general, Pasteur Institute, 25–28 rue du Docteur Roux, 75724 Paris, France. Mark Walport is director of the Wellcome Trust, Gibbs Building 215 Euston Road, London NW1 2BE, UK.

1. Murray, C. J., Lopez, A. D., Chin, B., Feehan, D. & Hill, K. H. *Lancet* **368**, 2211–2218 (2006).
2. Leroux-Roels, I. et al. *Lancet* **370**, 580–589 (2007).
3. Stephenson, I. et al. *J. Infect. Dis.* **191**, 1210–1215 (2005).

Competing financial interests: declared (see online article for details).

See Editorial, page 137. To discuss this article online visit <http://tinyurl.com/625swp>.

V. GHIRDA/AP

BOOKS & ARTS

Testing times for genomics

In a bid to popularize the science and ethics of human genetics, two new books fail to address developments in gene testing since completion of the Human Genome Project, says **Kathy Hudson**.

Two new books attempt to bring the issues raised by the Human Genome Project before a broader audience. In *DNA: Promise and Peril*, geneticists Linda and Edward McCabe try to be tour guides. In *Just Genes*, bioethics consultant Carol Isaacson Barash tries to teach. Ultimately, neither book goes far enough to improve the public's understanding, and neither addresses contemporary problems that bedevil policy-makers and the public.

Barash provides an interesting, if poorly edited, historical account of ethically troubling genetic policies and programmes. "The goal of the book is to improve your ability to identify and analyze ethical issues," she says, and opens well with a description of the ethical principles and theories that could be applied to the issues raised by genetics. But that is the last you hear of them. Subsequent chapters discuss the familiar subjects of genetic privacy, genetics and human behaviour, cloning and the like, yet fail to demonstrate how the theories apply. Instead, Barash provides long lists of questions, often more than a dozen, for each issue. Knowing what questions to ask is an essential first step, but a teacher should help the reader critically evaluate these issues.

In *Promise and Peril*, the McCabes fare only marginally better as tour guides. They wrote the book in part, they relate, so that "your concept of your own identity can truly be liberated". *Viva la liberación!* But reading *Promise and Peril* is like visiting Paris with a guidebook written in the 1930s: the landmarks are mostly the same, but the political and social landscape is utterly changed. It is possible to find the Eiffel Tower, but where are the Internet cafes, the supermarkets and the ethnic restaurants?

Rather than chart unfamiliar territory for their readers, the McCabes, and to a lesser extent Barash, rehash the arguments about genetic determinism made more eloquently by Dorothy Nelkin and M. Susan Lindee in *The DNA Mystique* (W. H. Freeman, 1995). The McCabes declare that there is a widespread "perception that individual human futures are



New genetic tests raise questions about privacy, but their reliability poses a bigger problem.

fully described in the sequences within each individual's genome" and set out to refute it. Barash similarly asserts that genes are commonly viewed as "encapsulating our innermost essence". Most readers would agree that a society that believes that traits, both medical and behavioural, are genetically hard-wired will ignore social policies and focus on genetic fixes. But how extensive is genetic determinism as a world view?

The McCabes argue that the notion that genes are not destiny has been "extremely difficult for some geneticists to accept".

This statement is puzzling. The tools of molecular biology and genetics, such as temperature-sensitive mutations, are rooted in the understanding that genes plus the environment equals phenotype. Disproportionate attention to genes arises not because scientists think that genes are everything, but because the genome is finite, knowable, and we have the tools to explore it. The complexity of the genome pales in comparison with that of human behaviour.

The favourite whipping boy in each book's

crusade against genetic determinism is the media. Yet serious studies of media representations of genomics fail to bear out their view, finding instead that determinism is seldom used by either the press or the public as a framing device for a discourse about genetics.

Both books devote considerable ink to privacy and genetic discrimination, especially misuse of genetic information by US health insurers and employers. Happily, the Genetic Information Nondiscrimination Act passed this year in the United States renders this discussion moot — the act will largely protect the privacy and integrity of genetic test results. Thus, the arguments forwarded by the authors can best be viewed as a historical footnote.

The most pressing policy question may not be the privacy of genetic test results, but the safety and accuracy of the tests themselves. This issue is at the heart of genomic medicine, yet is not addressed in either book. Barash mistakenly concludes that "deciding when pharmacogenetic testing is appropriate for commercial use will involve demonstrable proof of clinical validity and utility". In reality, genetic tests — including pharmacogenetic tests — can enter the marketplace and be sold directly to consumers without

Just Genes: The Ethics of Genetic Technologies

by Carol Isaacson Barash
Greenwood: 2007. 288 pp.
\$49.95, £27.95

DNA: Promise and Peril

by Linda L. McCabe and Edward R.B. McCabe
University of California Press:
2008. 356 pp. \$39.95, £23.95

any assessment of their usefulness or validity in a clinical setting.

The omission is notable given that Ed McCabe chaired the US Department of Health and Human Services Advisory Committee on Genetic Testing during its evaluation of genetic testing oversight. This committee concluded that government regulation was inadequate and put forth a series of policy recommendations. Because none of these recommendations was adopted, the issue still looms large — both for genetics and

geneticists, and for medicine more broadly.

Good guidebooks should point out new attractions as well as old, and new texts ought to cover contemporary issues as well as the historical. Neither the McCabes nor Barash accomplish this — both books are rooted in past discourses and ignore current issues of greater public health and policy urgency. ■

Kathy Hudson is director of the Genetics and Public Policy Center, Johns Hopkins University, 1717 Massachusetts Avenue NW, Washington DC 20036, USA.

contrast, organic chemistry is more than a century old, so the boundaries of patent claims are written clearly enough for competitors to understand what is claimed and what is not. Thus, companies can avoid infringement or resolve disputes quickly. Finally, the authors blame judges for expansionist tendencies that lead them to allow patents on “everything under the sun made by man”, resulting in a flood of biotechnology and software patents.

Bessen and Meurer use a lively, assertive style and are critical of patent-system stakeholders, including federal judges, the USPTO and patent lawyers. Throughout the book, they contrast the patent system’s rules about claim boundaries with the supposed success of real-estate law, a system that has clear, surveyed and publicly available land boundaries.

In the late 1990s, the authors contend, the US patent system failed publicly traded corporations. They estimate that the cost to such businesses of litigation over chemical and pharmaceutical patents was around \$4 billion in 2000, but this was more than outweighed by the profits associated with patents of about \$15 billion, a questionably low number given the value of blockbuster drugs. But for other industries, litigation costs were about \$12 billion and the profit was only about \$3 billion, also a questionably low number, and a net social loss. The authors recognize that owning patents always provides a competitive advantage for individual companies, even if the overall industry would be better off without them.

The authors offer several suggestions. Force inventors to clarify ambiguous claims with the USPTO. Convince judges to invalidate claims that are not definite and clear. Empower the

USPTO to give pre-litigation opinions on claim scope and infringement. Increase renewal fees to prune out unused patents. Establish a prior user defence so that independent discovery and use of a patented invention is not considered infringement.

Admirably, Bessen and Meurer reject what they call “faith-based policy” based on unsubstantiated beliefs about how patents help or hurt different companies and industries. Instead, they aim to provide the “first comprehensive empirical evaluation of the patent system’s performance”. They do well with studies of famous patent disputes, such as Kodak versus Polaroid (regarding instant

The cost of vague patents

Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk

by James Bessen and Michael J. Meurer
Princeton University Press: 2008. 352 pp.
\$29.95, £17.95

Adding fuel to the anti-patent fire, James Bessen and Michael Meurer argue in *Patent Failure* that US patents cost more than they should because their claims are too vague. Patent claims define the scope of a patented invention, and when they are unclear, competitors cannot determine whether or not they are infringing the patent and whether they should license or litigate.

Bessen and Meurer argue that ‘fuzzy’ claims for abstract inventions such as software, biotechnology and business methods make patents too expensive to society. Chemical and pharmaceutical patents, however, produce net economic benefits, as do patents obtained by small companies.

The authors spend most of their energy attacking software patents, describing skyrocketing litigation costs, increasing uncertainty, the thousands of patents on a single product, and other problems that hinder software innovation. They ask why the system works more poorly for software and biotechnology companies than for pharmaceuticals and chemicals. Their research leads the authors to posit several explanations. First, claims on small molecules are clearer and more comprehensible than the abstract claims in

biotechnology and software patents.

Second, the US Patent and Trademark Office (USPTO) and courts unwisely permit ‘premature’ patent claims on biotechnology and software products that have not and cannot yet be made. For example, Amgen’s patent on a method of making the red-blood-cell protein erythropoietin in hamster cells claimed all “non-naturally occurring” erythropoietin, including that made in human cells, even though erythropoietin had not yet been made that way.

Third, the courts have allowed patenting of early-stage biotechnology and software inventions that remain inchoate and abstract, increasing the likelihood of litigation against someone who independently develops an invention that falls within the scope of a patent claim. Fourth, patent law is in flux for the new fields of biotechnology and software. By



Microsoft’s legal chief Brad Smith (bottom left) and lawyers wait for a 2007 decision at the European Union court that fined the company over expensive patent licences.

J.-C. VERHAEGEN/AFP/GETTY IMAGES

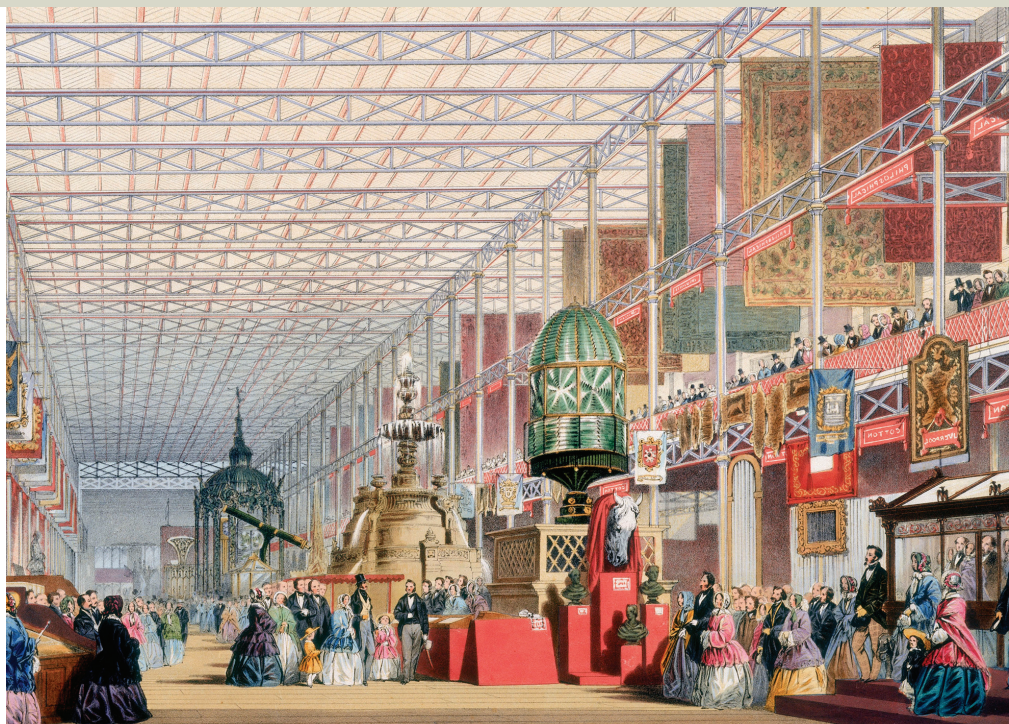
photography) and NTP versus Research in Motion (regarding the BlackBerry wireless device). Unfortunately, the book's strength — its reliance on data — is also a weakness. Most of the data are peripheral to traditional methods of patent valuation, such as market value, and date from the 1990s, so are only a trailing indicator of patent activity. The book's conclusions do not adequately reflect recent developments, including US court decisions that have narrowed patent scope, the open-source software movement and global pressures to limit software and medical patents.

Patent Failure focuses narrowly on financial value and costs of patents as private-property assets of publicly traded corporations. But patents serve other objectives that deserve greater attention. They are part of a dynamic, global intellectual-property system that drives innovation by balancing the exclusive rights of creators against the ability of others to access and copy creative works. The social benefits of the patent system, as it has evolved over the centuries, include promoting publication as an alternative to secrecy, rewarding investment in research, expanding international trade, allowing individual freedom and control over creative activities, promoting innovation races, helping innovators break into an established market allowing well-managed companies to supplant poorly managed competitors, and facilitating collaboration and technology transfer between public and private organizations. These benefits are undervalued or ignored in the book's financial calculations, and they contradict its conclusion of patent failure.

The book uses lively anecdotes and analogies, but this is ironic because the authors criticize other writers for relying on anecdotal information. And they overstate their case that patents would work better if they were more like real estate or copyright. The recent collapse in the real-estate market suggests that particular system is no panacea. And the authors' favourable view about the clarity of copyrights in the entertainment industry ignores frequent disputes over who owns rights to movies and music.

Patent Failure sets out in the right direction by comparing the corporate benefits of patents to their overall social costs, based on empirical data. And the book's conclusion may be sound, that the patent system would work better with clearer boundaries defined by patent claims. Ultimately, the book raises good questions, but leaves many unanswered. ■

Michael Gollin is a partner in the patent group at Venable LLP, 575 7th Street NW, Washington DC 20004, USA. He is author of *Driving Innovation: Intellectual Property Strategies for a Dynamic World*.



A new window on the Victorians

Victorian Glassworlds: Glass Culture and the Imagination 1830-1880

by Isobel Armstrong

Oxford University Press: 2008. 400 pp.
\$60.00, £30.00

Robert Lucas Chance was a man of unpredictable temper. The method of mass-producing artisan-blown cylinder glass pioneered at his factory near Birmingham, UK, in the 1830s was similarly prone to fiery eruptions, with coal firing difficult to control, and flames flashing unexpectedly out of the furnace when the molten glass was ready to be blown. Chance's name became synonymous with glass production in the nineteenth century, just as Hoover or Dyson are with vacuum cleaners today, and was peculiarly apt for the precarious nature of his business. The residual instability of the process, a contradictory amalgam of modern and archaic working practices, was made tangible in every sheet of cylinder glass by the traces of breath from individual blowers.

Chance's new processes nonetheless increased efficiency, facilitating the mass production of sheet glass and the manufacture of individual panes up to three metres square. The earlier Crown glass could reach a diameter of only half that, and was in a wasteful circular form. Historian Isobel Armstrong argues that Chance helped bring into being "an environment of mass transparency ... never before experienced".

The Victorian age embraced the new possibilities of cheap glass. The middle class developed a craze for popular microscopy, while Chance Brothers and Company glazed the shimmering, gargantuan Crystal Palace

(pictured) that housed London's Great Exhibition in 1851. The Victorians developed distinctive responses to this tantalizing material, from the bourgeois consumerism induced by the enticements within shop windows to the targeted window-breaking by working-class political activists. Victorian expansionism ensured that such responses extended across the oceans, safeguarded by lighthouses also glazed with Chance glass.

Victorian Glassworlds contends that technologies change experience. The widespread availability of full-length mirrors, particularly the fashionable cheval-glass mirror, altered how people perceived their bodies and prompted new ways of thinking about the nature of individual identity. London's Royal Coburg Theatre (now The Old Vic) installed a huge array of mirrors called a looking-glass curtain behind its stage in 1822. Audiences spent performances gleefully waving at themselves or ogling their superiors in more expensive seats, transforming themselves into performers and inducing an illusion of social unity.

Meanwhile, cylinder glass made possible the massive conservatories developed by John Claudius Loudon and Joseph Paxton in the 1830s and 1840s, such as Paxton's famous example at Chatsworth in Derbyshire. They allowed botanical specimens from across the expanding British empire to be grown despite the inclement weather, and forced more familiar flora to bloom out of season. Paxton's conservatory inspired his design for the Crystal Palace. Its success led him to propose placing much of central London under a ten-mile-long enclosed glass atrium known as the Great Victorian Way. Although it was never realized, this visionary scheme shows the potential for



FROM POLE ...

Acclaimed director Werner Herzog presents a personal view of Antarctica in his latest film, *Encounters At The End Of The World*. He meets some of the 1,000 scientists who work at McMurdo Station, the US research headquarters on Ross Island, and muses on the beauty of the blinding landscape. Underwater shots skim the bases of ice floes and follow researchers diving for new species (pictured). The film opened last month in New York, moves to selected US cinemas during the summer and will premiere at the Toronto International Film Festival in Canada in September.

<http://encountersfilm.com>

... TO POLES

A NASA telescope features in the documentary *BLAST!* Film-maker Paul Devlin followed his astrophysicist brother Mark's team for 5 years as they attempted flights of the Balloon-borne, Large-Aperture, Submillimeter Telescope (BLAST) in the Arctic and Antarctic. The researchers reflect on their science, philosophy and struggles to capture light from distant galaxies. The film opened in Toronto in April. It will show at film festivals worldwide, including Imagine Science in New York in October and Doc/Fest in Sheffield, UK, in November.

www.blastthemovie.com

SAVING LIVES

The use of retroviral drugs in treating AIDS is documented by eight noted photographers in *Access to Life*, on show at the Corcoran Gallery in Washington DC until 20 July. Supported by Magnum Photos and the Global Fund to Fight AIDS, Tuberculosis and Malaria, it chronicles how the drugs improve the lives of 30 patients worldwide, highlighting the importance of global funding for AIDS research. The exhibition will travel to Mexico City, Paris, London, Berlin and Rome.

www.corcoran.org

individual and social transformation ushered in by mass-produced glass.

Armstrong's assertion that technological changes were responsible, at least in part, for profound alterations in widely different areas of human experience is part of a trend in contemporary cultural history. It recalls Lynda Nead's *The Haunted Gallery* (Yale University Press, 2007), which argues that the invention of moving pictures at the end of the nineteenth century influenced other forms of visual art such as painting and sculpture, and on the very act of looking. The importance of material innovations has prompted historians and cultural critics to take a much greater interest in technological developments, and

offers a potential bridge between science and the humanities.

The sumptuous production quality and copious illustrations of *Victorian Glassworlds* make it seem at first glance just an attractive coffee-table book. In fact, Armstrong's compelling argument is complex and dense, with as much Freud and Foucault as flint or float glass. This challenging book is worth the effort, for the new window it opens on the crucially important Victorian period.

Gowan Dawson is senior lecturer in Victorian literature at the University of Leicester, University Road, Leicester LE1 7RH, UK. He is author of *Darwin, Literature and Victorian Respectability*.

The leaning tower of Perth

The Cosmology Gallery

Gravity Discovery Centre, Gingin, Australia

Opens 16 July 2008

A 40-metre-high tower juts from a vast expanse of bushland. Beside it sits a futuristic dome and a huge tripod, supporting a Foucault's pendulum that traces Earth's rotation. Signs invite visitors to "walk to the end of the solar system". This is the Gravity Discovery Centre, a public outreach facility at the development site of the Australian International Gravitational Observatory. Its Cosmology Gallery opens on 16 July.

The observatory, 80 kilometres north of Perth, Western Australia, is part of an international project to search for gravity waves — moving ripples in space-time. Gravity waves were predicted by Einstein in his general theory of relativity, but have yet to be directly detected.

The new educational centre is the brainchild of scientists at the University of Western Australia. It is modelled on three buildings in Pisa, Italy: the cathedral, the baptistry and the leaning tower, where Galileo carried out his famous experiment on gravity. The tower is a triangular prism of steel that leans by 15° from the vertical, farther than its namesake in Pisa, and is touted as the most deliberately leaning built structure in the world. Students climb to the top to repeat Galileo's observation that all objects fall at the same rate owing to gravity. The main visitor centre, containing exhibits about gravity and the cosmos, represents the cathedral.

The Cosmology Gallery echoes Pisa's baptistry and mixes science, art, culture and religion. Like the molecular structure of a carbon-60 buckyball, it is composed of interlocking hexagons and pentagons. The tiled floor was designed



by mathematician Roger Penrose. A walkway spirals up the inside of the gallery towards higher levels and a timeline of the Universe.

Commissioned artworks explore our place in the Universe from many perspectives, including Christian, Buddhist, Hindu, Islamic and Australian Aboriginal. They include a Buddhist mandala and photographs showing re-enactments of Balinese creation legends. *Bibbulmun Religion* (detail, pictured), a painting by Aboriginal artist Toogarr Morrison, incorporates traditional motifs and modern images of spiral galaxies. Other paintings depict the phases of the Moon and the southern sky at night.

Just as gravity waves provide new ways to study the Universe, so the juxtaposition of scientific and cultural traditions in the gallery challenges us to broaden our thinking about cosmology.

Keith Hall is a chemist and freelance writer based in Perth, Western Australia.

ESSAY



D. PARKINS

Behind the looking-glass

To understand how mirror neurons help to interpret actions, we must delve into the networks in which these cells sit, say **Antonio Damasio** and **Kaspar Meyer**.

A remarkable discovery was made more than ten years ago: some neurons in the brains of macaques are active both when a monkey moves and when it sees a person move in a comparable way¹. The lead researchers, Giacomo Rizzolatti and Vittorio Gallese, called the cells involved 'mirror neurons'. This evocative name, and the significant implications of the finding, led to a surge of scientific and public interest in these cells. But perhaps the name was too evocative for the finding's own good. It seems to have tempted people into thinking of these neurons as tiny, miraculous mirrors that allow us to understand each other, diverting attention from the search for how they work.

After mirror neurons were found in two main regions of macaque brains^{1,2}, corresponding results were found in human brain scans³. Researchers began to invoke mirror neurons to explain the fundamental task of

how humans and other primates recognize the actions and, by extension, the intentions of others. Mirror-neuron activity is thought to generate a more-or-less explicit simulation of others' actions in the observer's brain. If this is the case, then simulation, rather than inference, would provide a deep understanding of those movements. Some have proposed that mirror neurons are a vital part of how

"Hearing a peanut being broken without seeing or feeling it triggers a whole neural network."

infants first learn to imitate other people's actions; of why seeing a grimace can trigger empathetic feeling in a watcher; and even of how language is learned. Although simulation is unlikely to account entirely for how people understand what they see other people do, the identification of mirror neurons was an important step on the path to unpacking that process.

Unfortunately, the lack of a satisfactory explanation for how the simulation process occurs has allowed the notion to take hold

— among the public and scientists alike — that mirror neurons accomplish mirroring on their own in some mysterious way.

Surely, if 'action understanding' is to occur by internal simulation, the process must enlist both motor and sensory systems in the brain. Mirror neurons cannot accomplish this task alone. Rather, they must be acted on by other structures; and, just as importantly but a fact generally neglected, mirror neurons must act on other structures.

Mutual support

Unravelling how mirror neurons work requires knowledge of the complex architecture in which these cells are embedded. A model of neural architecture proposed by one of us (A.D.)⁴ nearly 20 years ago may help with that. Impressive findings from mirror-neuron research at the single-cell level have added persuasive empirical support for this early model.

The model, originally called 'time-locked multimodal retroactivation', was derived

during the 1980s from observations in patients with brain lesions. Some patients with damage in the anterior (higher-order) sectors of the temporal cortex, for example, could not recall complex memories that combine separate components of a specific event, such as a face, a name, an action or a place. These patients could not recognize close friends or relatives, nor could they picture unique events in time and place, such as their own wedding or the birth of a child. Yet they could easily imagine or recognize a picture of a non-specific wedding or a baby.

In other words, anterior damage did not preclude their retrieval of mental representations of objects, places or actions; but it did stop them recalling certain combinations of representations that signified a particular person or event. Only damage to posterior sectors of the cerebral cortex, near sensory and motor cortices, impaired access to separable memory components. Given what was already known about memory systems, this suggested that anterior sites held the key to some process needed to reconstruct, elsewhere, the parts that made up a complex memory.

Mind mapping

An idea for a brain architecture that could explain these findings was inspired by experimental neuroanatomy studies showing that, surprisingly, signals are conveyed within the brain in both forwards and backwards directions. For example, signals are obviously sent from the eye to the visual cortex and on to areas of higher-level processing in the brain. But these high-level areas also send signals back to the visual cortex, and even to the visual thalamus, below the level of the cortex⁵. The same forward-backward signalling arrangement was found for the hippocampus — a brain region involved in making factual memories⁶.

All this led to a model that posited the existence of 'convergence-divergence zones' (CDZs). These neural ensembles collect signals from separate sites, and signal back to those sites. When several signals converge on a CDZ, the ensemble creates an abstract record of the coincident activations — a memory trace, in other words. The model contained two broad types of CDZ. 'Local CDZs' were proposed to coordinate information within regions close to a sensory cortex, such as the visual cortex. These local hubs were proposed to converge on 'non-local CDZs' in higher-order sectors of the brain.

In this view, when a monkey breaks open a peanut⁷, local CDZs collect information about various sensory inputs, and feed these to a non-local CDZ that records the coincident

information about the sound, sight and feel of this action. The CDZ does not hold all the details of this information; rather, it contains the potential to retroactivate the separate auditory, visual, tactile and motor sites, and thus reconstitute the original distributed set of memories and information. Imagine the CDZ as a repository of instructions for a book that must be printed by several different printers. Having the instructions alone will not get you the book, and the printers alone will not help either. You need both to get the final product.

In future, hearing a peanut being broken without seeing or feeling it triggers a series of events. First, it activates the auditory cortices and local auditory CDZ; second, it activates the non-local CDZ that previously collected the memory trace associated with this sound; third, it precipitates simultaneous signalling outwards from this non-local CDZ to all the local CDZs involved in the original event (motor, visual, auditory); fourth, it reactivates all or some of these sites. This leads to a more-or-less successful replay of the coincident set of separate brain activities that accompanied the monkey breaking open a peanut.

Looked at in this way, mirror neurons correspond to non-local CDZs. Their connections to other CDZs, and their ability to collect and distribute signals based on learned experience, allow the brain to reconstruct an action from only part of the story. A whole neural network underlies the understanding of action, rather than a single anatomical site or even a single cell. The monkey's comprehension of the sound of a cracking nut is not created just by mirror-neuron sites, but also by the nearly simultaneous triggering of widespread memories throughout the brain.

The neurons at the heart of this process, and at the heart of non-local CDZs, are not so much like mirrors, after all. They are more like puppet masters, pulling the strings of various memories.

Recent findings are in line with this view. Studies in humans and monkeys show that the neural network stimulated by watching an action goes beyond the original mirror-neuron sites; it encompasses more widespread sensorimotor cortices^{3,8}. Conversely, carrying out an action recruits sensory cortical areas even when subjects can neither see nor hear the actions they perform⁹. This lends further support to the notion that the neural description of an action goes far beyond its

motor components. At least one other study has invoked the necessity of convergent signals into mirror-neuron areas to explain such observations¹⁰. The CDZ model, and our interpretation of mirror neurons, adds the aspect of divergence.

Neural networks

The CDZ framework allows us to see the role of mirror neurons more clearly. Cells in mirror-neuron areas do not themselves hold meaning, and they alone cannot carry out the internal simulation of an action. This runs counter to the understandable public misconception that mirror neurons alone 'mirror' action. Rather, mirror neurons induce widespread neural activity based on learned pat-

terns of connectivity; these patterns generate internal simulation and establish the meaning of actions. Mirror neurons pull the strings, but the puppet itself is made of a large brain network.

The CDZ model was well received, but the lack of single-cell experiments providing direct evidence for this functional architecture has limited the application of the idea. The mirror-neuron evidence sits well alongside the model, and each seems to make more sense in light of the other. But this is not proof. The ultimate test of the convergence-divergence model, and its explanation of how mirror neurons do what they do, depends on the ability to record brain activity simultaneously from neurons in separate sites, and on probing the underlying connectivity between neural areas. Such goals are within reach, albeit technically difficult to achieve. In the meantime, bringing together mirror-neuron data and the CDZ model could guide future efforts to explain the relationship between what we see and what we do. ■

Antonio Damasio and Kaspar Meyer are neuroscientists at the Brain and Creativity Institute of the University of Southern California, Los Angeles, USA.

"The neurons are not so much like mirrors, after all. They are more like puppet masters, pulling the strings of various memories."

1. Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. *Brain* **119**, 593-609 (1996).
2. Fogassi, L., Gallese, V., Fadiga, L. & Rizzolatti, G. *Soc. Neurosci.* **24**, abstr. 257.5 (1998).
3. Buccino, G. et al. *Eur. J. Neurosci.* **13**, 400-404 (2001).
4. Damasio, A. R. *Cognition* **33**, 25-62 (1989).
5. Rockland, K. S. & Pandya, D. N. *Brain Res.* **212**, 249-270 (1981).
6. Van Hoesen, G. W. *Trends Neurosci.* **5**, 345-350 (1982).
7. Kohler, E. et al. *Science* **297**, 846-848 (2002).
8. Raos, V., Evangelou, M. N. & Savaki, H. E. *J. Neurosci.* **27**, 12675-12683 (2007).
9. Gazzola, V., Aziz-Zadeh, L. & Keysers, C. *Curr. Biol.* **16**, 1824-1829 (2006).
10. Keysers, C. & Perrett, D. I. *Trends Cogn. Sci.* **8**, 501-507 (2004).

NEWS & VIEWS

PALAEOLOGY

Squint of the fossil flatfish

Philippe Janvier

Evolutionary biologists have floundered when trying to explain how the asymmetrical head of flatfishes came about. 'Gradually' is the answer arising from exquisite studies of 45-million-year-old fossil specimens.

Minor individual variations apart, external asymmetry is rare in vertebrates. It is limited to hagfishes, which possess an enigmatic pore (possibly a modified gill pore) on their left side; to some cetaceans, whose asymmetrical tusks may be linked to certain feeding behaviours that favour the right side; and to the celebrated example of flatfishes (such as turbot, flounder, halibut, sole). On page 209 of this issue¹, Friedman adds a contemporary chapter to the long history of thinking about flatfish evolution.

The asymmetry of adult flatfishes is remarkable because the young are perfectly symmetrical. The entire head structure is rapidly remodelled during metamorphosis, involving the migration of one eye to the opposite side, close to the other eye. The polarity of this migration varies from one species to another, and even within the same species; some species generally have both eyes on either the left side (turbot) or the right side (halibut, sole), but all living and fossil flatfishes known to date show an 'eyed' side and a 'blind' side (Fig. 1). Only in the most primitive living flatfish, the 'spiny turbot' (*Psettodes*), is one eye (either the left or right one, with equal frequency) situated close to the dorsal margin of the head, although it is still on the same side as the other eye.

The lack of any intermediate condition in living and fossil adult flatfish species has raised endless debates and conjectures about the origin of this striking and obviously derived character (that is, one modified from the primitive symmetrical body plan). Lamarck, Darwin, Mivart, Goldschmidt and other great figures of early evolutionary biology all dealt with the question, coming up with different answers and evolutionary scenarios that generally put the case either for gradual evolution or for evolution by leaps ('saltation'). Friedman now takes the debate into the twenty-first century. He provides unambiguous evidence that two adult forms of primitive fossil flatfishes display incipient — but incomplete — head asymmetry. This discovery shows that the evolution of flatfish asymmetry was indeed gradual, and so will help lay to rest a favoured creationist example.

The fossils that Friedman¹ has studied are



Figure 1 | Eye eye. During the development of extant flatfishes, such as this plaice (*Pleuronectes platessa*), one eye has migrated round the head to lie on the same side as the other. So these fishes have an 'eyed' (up) side and a 'blind' (down) side suitable for their bottom-dwelling lifestyle.

about 45 million years old and are members of two genera, *Heteronectes* (newly described by Friedman himself) and *Amphistium* (which has been known for a long time). The specimens all come from the exceptional fossil-fish locality of Bolca, Italy, and were once regarded as possible flatfish relatives on the basis of their postcranial skeleton. However, this suggestion was later dismissed because their skull asymmetry was unclear and was attributed to distortion during fossilization. None of the fossils had been prepared or studied with the methods now used by fossil-fish specialists, in part because the fossils were regarded as too precious. Some of them are of historical interest. For example, those in a Paris museum were 'donated' by Count Giovanni Battista Gazzola of Verona to Napoleon Bonaparte after his collection had been seized by French troops² (although Gazzola subsequently received seeds from Paris in exchange).

Friedman¹ was allowed to treat a single specimen housed in the Naturhistorisches Museum, Vienna, with weak acid, and to carry

out computed tomography imaging of the skulls of two specimens housed in the Natural History Museum in London. This provided him with unprecedented anatomical information, from which he could rule out distortion as an artefact of fossilization. The skull of these fishes is clearly asymmetrical, with one eye looking up on one side and the other down on the opposite side (Fig. 2, overleaf), possibly at equal frequency as in the living *Psettodes*.

These fossil fishes are coeval with the earliest, typically asymmetrical fossil soles, and must have had an earlier and rather long unrecorded history, during which they survived perfectly well despite their squinting eyes. What could have been the selective advantage of this unusual anatomy, relative to that of the presumed closest symmetrical relatives of flatfishes such as batfishes, sea sweeps and pompanos? Many possible scenarios have been proposed to account for the propensity of flatfishes to lie on one side or the other and for eye migration³. Autenrieth⁴ explained eye migration by simply concluding that the eye facing

K. TELNES/IMAGEQUESTMARINE.COM

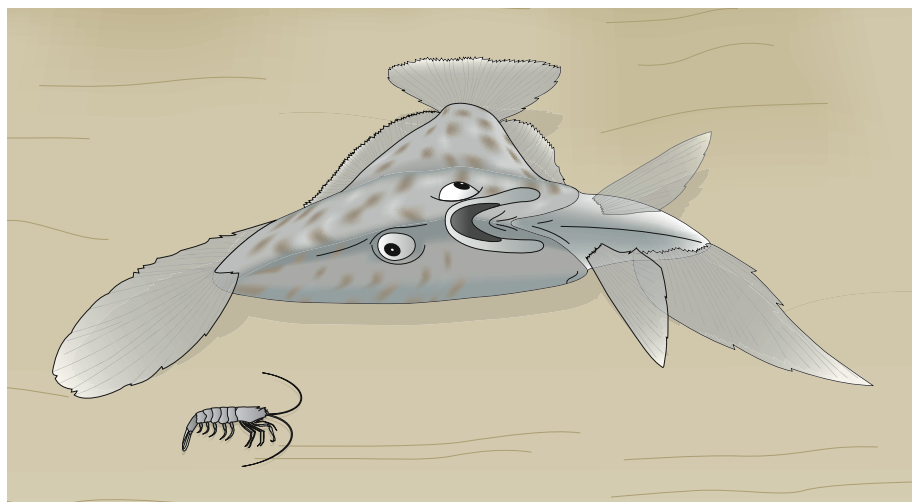


Figure 2 | The benefit of squinting? Living flatfishes sometimes raise their head above the sea bottom by bending down their fin rays. Ancestral flatfishes, such as the 45-million-year-old *Amphistium* depicted here, already possessed an asymmetrical skull but — as Friedman¹ describes — retained eyes on either side of the head. As shown in this drawing, they may have used the strategy of lifting their head to look for prey with one eye and for predators with the other. For ease of visual orientation, rotate the image by 90° clockwise. (From a sketch by P. Janvier.)

down would have been useless. Lamarck⁵ invoked the need for flatfish precursors to survive in very shallow water. In the same vein, Malm⁶ cited the propensity of the symmetrical larval flatfishes to rest on one side, but this was later dismissed. Darwin⁷ pointed out the correlation between eye migration and jaw asymmetry, which brought the prey within closer reach of the mouth⁵. More recent scenarios also invoke internal asymmetry, which

generates a lateral tilt at rest in many deep-bodied fishes, including batfishes.

Friedman¹ himself points out that some modern flatfishes, when lying in ambush, tend to lift up their head from the ground by bending down their dorsal- and anal-fin rays. Similar behaviour in the flatfish ancestors described here may have been the rule, thereby allowing both eyes to be functional even though they lay on either side of the head. Perhaps one eye

was used to watch for predators and the other to search for prey (Fig. 2). Squinting may have had a selective advantage for a while, but regaining binocular vision — albeit on one side — ensured the lasting success of flatfishes.

One of the informative properties of fossils is that they may display character combinations that no longer exist in the living world, and so provide 'intermediate' or 'transitory' forms or organs that support some evolutionary scenario or other. On occasion, this property has been pooh-poohed, even by palaeontologists who have considered that fossils can help in refining relationships already inferred from living species, but tell us little about the process of evolutionary character transformation⁸. In the case of the fossils described by Friedman, however, one cannot but admire the vindication of a prediction, made by Darwin, of a gradual eye migration during flatfish evolution that mirrors the metamorphosis of the living forms. ■ Philippe Janvier is at the Muséum National d'Histoire Naturelle, UMR 5143, CNRS, 8 Rue Buffon, Paris 75005, France.

e-mail: janvier@mnhn.fr

1. Friedman, M. *Nature* **454**, 209–212 (2008).
2. Faujas de Saint-Fond, B. *Essai de Géologie* (Dufour, 1809).
3. Chanet, B. & Chapleau, F. *Bull. Soc. Sci. Nat. Ouest Fr.* **27**, 101–108 (2005).
4. Autenrieth, J. H. F. in *Arkiv für Zoologie und Zootomie* Vol. 1 (ed. Wiedermann, C. R. W.) 47 (1800).
5. Lamarck, J. B. *Philosophie Zoologique* (Dentu, 1809).
6. Malm, A. W. K. *Svenska Vetensk-Akad Handl.* **7**(4), 1–28 (1868).
7. Darwin, C. *On the Origin of Species by Means of Natural Selection* 6th edn (Murray, 1872).
8. Patterson, C. *Annu. Rev. Ecol. Syst.* **12**, 195–223 (1981).

PLANETARY SCIENCE

The early Moon was rich in water

Marc Chaussidon

Analyses of lunar volcanic glasses show that they are rich in volatile elements and water. If parts of the lunar mantle contain as much water as Earth's, does this imply that the water has a common origin?

The Moon's chemical composition differs from Earth's. It is enriched by a factor of two to three in refractory elements (those that condense first from a high-temperature gas) such as aluminium, calcium and titanium; most easily vaporized (that is, volatile) elements, such as sodium and potassium, are rare; and it is considered to be almost devoid of water¹. The partial loss of all but the least volatile elements is generally thought to be a signature of the lunar forming event. High temperatures were reached in the debris ejected when a Mars-sized body hit the 30-million-year-old proto-Earth. On page 192 of this issue, Saal *et al.*² present data counter to this classic image of lunar geochemistry. They show that some parts of the lunar mantle may contain a few hundred or even a few thousand

parts per million (p.p.m.) of water.

The water content of Earth's upper mantle can be estimated from the analysis of mid-ocean-ridge basalt (MORB) glasses — more precisely, from the analysis of micrometre-sized melt inclusions that were trapped within crystals of the mineral olivine when these crystals grew deep in their parent mantle melts³. Water is concentrated in melts during mantle melting, so calculating the composition of the original solid source material requires corrections for, among other things, the distribution coefficient of water in the melt and the amount of melting undergone. From the concentration ratios of water to the element cerium⁴, and the absolute water contents of MORB glasses and their melt inclusions (400–1,600 p.p.m.^{3,4}), the water content of Earth's upper mantle has been

calculated to be about 150 p.p.m. The water content of the bulk Earth (mantle plus surface) is estimated to be 350 ± 50 p.p.m.⁵.

Saal *et al.*² have taken a similar approach using samples of lunar soils collected during the Apollo missions (Fig. 1). The soils contain glasses that colour the soils green (collected by Apollo 15) or orange (collected by Apollos 11 and 17), and are probably of volcanic origin. These glasses have peculiar compositions: the green glasses are rich in magnesium, the orange ones rich in titanium. Most glasses found on the Moon were produced by the melting of lunar rocks during meteorite impacts. But the green and orange glasses lack the typical characteristics of impact glasses⁶, such as high levels of nickel derived from the incoming meteorite, or fragments from the impacted rocks. In addition, the ages of the green and orange glasses are similar to those of the surrounding lunar basaltic rocks, which links them to the geological evolution of the Moon. These green and orange glasses, (spherules 0.1–0.4 mm in size) are believed to be the products of volcanic eruptions on the surface of the Moon — 'fire fountaining' events that are well known for terrestrial volcanoes.

Coating the surfaces of the green and orange glasses are a variety of volatile elements¹. These

NASA



Figure 1 | Glass bead game. Orange soil discovered by the Apollo 17 astronauts on the rim of the Shorty crater on the Moon. The colour is due to the presence of numerous glassy spherules, 0.1–0.4 mm in diameter, which are orange because of their high titanium content. The glasses are thought to be products of volcanic fire fountains. They were buried in the soil after eruption and exposed during the impact responsible for the crater. The high concentrations of water found within the glasses by Saal *et al.*² suggest that the Moon may have had a damper origin than previously thought.

volatiles are assumed to have condensed onto the glasses from hot, transient plumes of gas associated either with lunar volcanism or meteorite impacts. However, Saal *et al.*² have looked for the presence of indigenous volatiles deep within the glasses. Using a technique known as secondary ion mass spectrometry (ion microprobe), the authors found significant amounts of chlorine, fluorine, sulphur and water, with the highest concentrations at the cores of the glass beads. This relationship is the opposite to that which would be seen if the volatiles had been added to the glass by any process, including contamination back on Earth, after its formation. The volatiles are present at depths of between 18 and 140 μm (and presumably deeper in large enough fragments). This also precludes a significant contribution of solar wind, because solar ions penetrate to a depth of only about 0.1 μm , even if hydrogen from the solar wind can diffuse to greater depths after its implantation⁷.

The simplest explanation is that these volatiles were present in the molten glass when it formed. Their distribution profile would then be a result of volatile loss by diffusion from the surface of the glasses during the short period when solidifying droplets were exposed to the lunar vacuum.

Modelling the initial water content before such degassing depends on the size of the

droplets, the diffusion coefficient of the water in the molten glass, and the temperature path encountered between eruption and quenching (the period during which volatiles could be lost). Saal *et al.*² calculate that the observed concentration profiles result from high initial levels of volatiles, although the precise concentrations are uncertain. Initial water contents could have been as low as 260 p.p.m. or as high as several thousand parts per million, with best fits to the data around 750 p.p.m., values similar to those in the MORB glasses. Most of the variations between samples probably resulted from variations in the degassing process.

These results raise many questions. Are the volatile contents of the melts that formed the green and orange glasses typical for the Moon? Can the general scarcity of most volatile elements on the Moon be reconciled with the apparent abundance of sulphur, chlorine, fluorine and especially water in the lunar glasses? What happened to all the water during the Moon's formation? And if the Moon is not bone dry, where did the water come from?

Another mystery in lunar geochemistry is that the relative abundance of the various oxygen isotopes on the Moon is the same as that on Earth to within 0.0005% (ref. 8). Different explanations have been proposed for this puzzling observation (Earth, Mars and all meteorites have different oxygen isotopic



50 YEARS AGO

In the past ten years the number of television licences in Britain has grown from less than 15,000 to nearly eight million, and the estimated number of adults aged sixteen or more in homes with television sets from 80,000 (all in the London region), which was 0.2 per cent of the total adults, to 21,850,000 or 57.9 per cent. For every person who had a television set in 1947 there are five hundred to-day, and six out of ten adults can see television in their own homes ... The United Kingdom now ranks third with Canada in its ratio of persons to television receivers (6); only the United States (4) and Hawaii (5) exceed it ... Published evidence as to the effect of television on social habits and hobbies does not indicate a wide difference between the habits of those owning television sets and those without, in many activities such as sport, gardening, theatre-going and card-playing: the most significant differences appear in reading, attendance at the cinema, church-going and dancing.

From *Nature* 12 July 1958.

100 YEARS AGO

Count Zeppelin last week made a remarkably successful flight in his new airship ... The distance covered is estimated at 250 miles, and the journey lasted twelve hours. The greatest height reached by the airship's own engine-power is stated to be some 750 metres, and the highest speed 15.3 metres per second ... We notice that Count Zeppelin has received a telegram of congratulation from the German Emperor.

ALSO:

The prize of 10,000 francs (400l.) offered by M. Armengaud to the first *aéroplane* to remain in the air for a quarter of an hour was won by Mr. Farman on Monday ... Mr. Farman made a flight with his apparatus which lasted 20m. 20s. according to official timing. He covered a distance of about eleven miles.

From *Nature* 9 July 1908.

50 & 100 YEARS AGO

compositions). One theory is that turbulent exchange of volatile material between Earth and a slowly coalescing Moon, for a few hundred years after the impact that formed it, caused oxygen isotopic equilibration between the two bodies⁹. Such a process could also have transferred volatiles such as water from Earth to the Moon.

This might be one explanation for the high water contents found in the green and orange glasses, but it is not the only one. The ratio of deuterium to hydrogen found on Earth is consistent with that expected if all terrestrial water has been delivered by chondritic meteorites after the cataclysm that formed the Moon had driven off most of the hydrogen present previously⁵. Is a late arrival of volatiles also possible for the Moon? Clearly, the next step is to measure the deuterium to hydrogen ratios in the green and orange glasses, not least because the only such data currently available for an orange

soil¹⁰ are probably marred by contamination with water after the samples were brought to Earth.

Marc Chaussidon is at CRPG, Nancy Université, INSU/CNRS, UPR 2300, BP 20, 54501 Vandoeuvre-lès-Nancy, France.
e-mail: chocho@crpg.cnrs-nancy.fr

1. Taylor, S. R. *Planetary Science: A Lunar Perspective* (Lunar & Planetary Inst., Houston, 1982).
2. Saal, A. E. *et al. Nature* **454**, 192–195 (2008).
3. Sobolev, A. V. & Chaussidon, M. *Earth Planet. Sci. Lett.* **137**, 45–55 (1996).
4. Saal, A. E., Hauri, E. H., Langmuir, C. H. & Perfit, M. R. *Nature* **419**, 451–455 (2002).
5. Marty, B. & Yokochi, R. *Rev. Miner. Geochem.* **62**, 421–450 (2006).
6. Delano, J. W. *J. Geophys. Res.* **91**(B4), D201–D213 (1986).
7. Hashizume, K., Chaussidon, M., Marty, B. & Robert, F. *Science* **290**, 1142–1145 (2000).
8. Wiechert, U. *et al. Science* **294**, 345–348 (2001).
9. Pahlevan, K. & Stevenson, D. J. *Earth Planet. Sci. Lett.* **262**, 438–449 (2007).
10. Epstein, S. & Taylor, H. P. *Geochim. Cosmochim. Acta* **2** (Suppl. 4), 1559–1575 (1973).

CANCER

An unexpected addiction

John D. Shaughnessy

Both oncogenes and normal genes can mediate the development and progress of cancer. What used to separate their effects was cancer's dependence on, or 'addiction' to, oncogenes but not normal genes. Not any more.

A central tenet of cancer is that it arises from a single cell through progressive acquisition of genetic mutations that lead to the activation of oncogenes (cancer-promoting genes) and/or inactivation of tumour-suppressor genes. Cancer's 'addiction' to oncogenes is sometimes so strong that even brief inactivation of a single oncogene can cause remission in model systems^{1,2}, implying that oncogenes are the 'Achilles' heel' of cancers³. On page 226 of this issue, Shaffer and colleagues⁴ reveal that the cancer multiple myeloma is similarly addicted to a non-oncogene, the product of which — the IRF4 transcription factor — has a central role in health.

Myeloma is a cancer of immune cells, specifically of antibody-secreting plasma cells in bone marrow. These cells arise from B cells, which, when stimulated by antigens, show increased activity of IRF4, leading to their differentiation into plasma cells⁵ (Fig. 1a). Multiple myeloma covers a spectrum of distinct molecular subtypes, and the disease is driven by several oncogenes^{6,7}. Chromosomal abnormalities such as translocations are common in multiple myeloma, and the bone-marrow microenvironment provides support for tumour-cell survival, mainly through growth factors secreted by bone-marrow stromal cells. These cells, and the bone-marrow microenvironment in general, can also protect myeloma cells from being killed

by chemotherapy or cytokine deprivation.

Shaffer *et al.*⁴ wanted to know what genes are crucial for the proliferation and/or survival of myeloma cells. They screened about 10% of all known genes, reducing their expression by the technique of RNA interference (RNAi) using small hairpin RNAs (shRNAs). Several oncogenes, such as *FGFR3*, *MMSET*, *MAF*, *MAFB* and the three *cyclin D* genes, are activated by chromosomal translocations associated with myeloma, to initiate myeloma formation⁸. But Shaffer and colleagues' screen did not identify an effect of suppressing these oncogenes, suggesting that the survival of myeloma cells does not depend on them. Instead, they found that, although the *IRF4* gene is not mutated in myeloma, depletion of IRF4 protein rapidly kills all myeloma cells, but has little effect on cells of another type of blood cancer, lymphoma.

That the inappropriate activation of a normal protein can mediate cancer-cell proliferation and survival is well documented. Also, previous studies^{6,9} have found that IRF4 is highly active in both normal plasma cells and primary myeloma. What is surprising about Shaffer and colleagues' observations is that, despite its battery of oncogenic mutations, myeloma is addicted to a normal gene.

What is the effect of IRF4 malfunction in myeloma? Gene-expression profiling studies¹⁰ have shown that a subset of myeloma exhibits features both of normal plasma cells

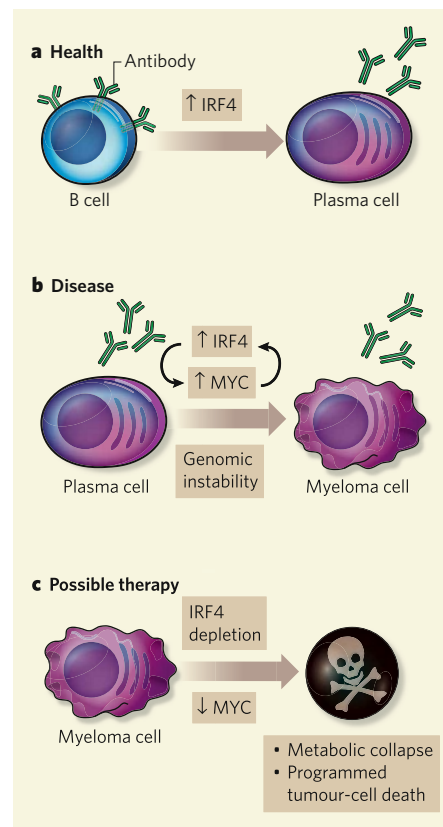


Figure 1 | IRF4 in health and disease. **a**, Normally, antigen-stimulated B cells differentiate into antibody-secreting plasma cells, a process that is partly driven by increased expression of the transcription factor IRF4. **b**, Like their normal counterparts, myeloma cells also express IRF4. Schaffer *et al.*⁴ find that, although *IRF4* is not an oncogene, it has a unique transcriptional repertoire of target genes in myeloma cells related to both its targets in plasma cells and in antigen-stimulated B cells. IRF4 also regulates the expression of the MYC protein, which itself regulates IRF4 through an autoregulatory loop. Apart from these IRF4-mediated effects in myeloma cells, genomic instability and both oncogenic mutations that initiate tumorigenesis (primary mutations) and secondary oncogenic mutations also contribute to myeloma. **c**, Despite myeloma cells harbouring oncogenic mutations, depleting IRF4 in these cells causes decreased expression of numerous genes including the *MYC* oncogene, as well as metabolic collapse and programmed cell death. These results suggest that inhibiting IRF4 might represent a new treatment for myeloma.

and of their precursors — antigen-stimulated mature B cells. Shaffer *et al.* propose that IRF4 might direct the gene-expression program that results in this abnormal feature of myeloma cells. One of their main findings is that, in myeloma, the *MYC* gene (which when mutated can potentially become an oncogene) is a direct target of IRF4 (Fig. 1b).

Expression of *IRF4* increases in both plasma cells and myeloma cells, but that of *MYC* is higher only in myeloma cells⁹. *MYC* has a central role in the formation of both myeloma and two other plasma-cell tumours,

plasmacytomas and Burkitt's lymphoma¹¹. The authors therefore investigated the relationship between MYC and IRF4 expression. They find that, in myeloma cells, these two proteins mutually reinforce each other's expression. IRF4 seems to be a master regulator, as its direct targets — many of which are also MYC targets — include mediators of metabolic control, membrane biogenesis, cell cycle, cell death, regulation of transcription, plasma-cell differentiation, glucose metabolism and the production of the cell's energy molecule ATP. It is therefore plausible that, on IRF4 depletion, 'metabolic collapse' also contributes to myeloma-cell death.

As myeloma is a slow-growing tumour, it has been proposed¹² that myeloma cells are differentiated descendants of transformed B cells, which would be the myeloma progenitor cells. IRF4 is inactive in B cells, and IRF4-targeted shRNA does not kill these cells. So treatment of myeloma with IRF4 shRNA, which will kill the shRNA-sensitive plasma cells, should supposedly reveal those rare cancerous myeloma progenitors that are resistant to IRF4 shRNA. But despite extensive investigations, Shaffer *et al.* did not observe such a phenomenon. This suggests that if a 'myeloma stem cell' does exist, it must reside at a stage in B-cell differentiation that requires IRF4.

Despite remarkable improvements in the outcome of myeloma treatment over the past decade, cures are elusive and the clinical management of this cancer remains challenging¹³. The non-uniform and disseminated nature of myeloma, and an inherent and acquired drug resistance resulting from interactions of tumour cells with the bone-marrow microenvironment¹⁴, worsens the prospect of finding new treatments. So the fact that a non-oncogene might be a new therapeutic target for nearly all forms of myeloma is a welcome advance (Fig. 1c). But before contemplating whether Shaffer and colleagues' observations can be translated into the clinic, we must consider whether these data themselves have one or two weaknesses.

The authors performed their studies in myeloma cell lines. All myeloma cell lines exhibit genetic features of highly aggressive disease, which are observed in only 15–20% of newly diagnosed myelomas⁷. Moreover, patients with such molecularly defined high-risk features are not benefiting from current treatments. So, although the finding that depleting IRF4 can kill cells with high-risk genetic features is exciting, one wonders whether IRF4 depletion has lethal effects only in the presence of these high-risk genetic features.

Similarly, the multiple myeloma subtype known as hyperdiploid disease, which makes up nearly 50% of primary myeloma cases, is not represented in the myeloma cell lines Shaffer *et al.* studied. It is therefore unknown whether IRF4 depletion will have similar effects on this and other myeloma subtypes. The use of cell lines leaves yet another question open: can the

bone-marrow microenvironment overcome the lethal effects of IRF4 loss on myeloma cells?

With Shaffer *et al.* testing only about 10% of human genes in their RNAi screen, is it possible that other genes exist that might be more amenable to therapeutic manipulation and whose depletion will have an effect similar to IRF4 loss? Cancer cells are notorious for developing resistance to chemotherapy; will the same hold true for the effect of shRNAs on IRF4?

The authors⁴ took advantage of extensive genome-wide expression-profiling data of primary myeloma disease^{7,11}. This publicly available myeloma gene-expression atlas¹⁵ provides a comprehensive genomic landscape of the cancer relative to a normal genome. Myeloma cells express about a third of their entire genome. An RNAi library designed to specifically target genes expressed only in myeloma might be a useful step in investigating their individual contribution to cancer and their potential as therapeutic targets. It would be equally important to take the converse approach and screen tumours for the lethal effects of introducing genes whose expression is lost in myeloma cells.

With patient-specific, tailored therapies as a central goal in the post-genomic era, is it

possible that more effective shRNAs or drugs emerging from them — easier to deliver and less likely to have adverse side effects — will be identified that target specific genes affected in particular subtypes of multiple myeloma? But perhaps the most crucial question of all is whether non-oncogene additions are found in other cancer types, especially tumours of epithelial origin, which represent the bulk of human cancers. ■

John D. Shaughnessy is at the Myeloma Institute for Research and Therapy, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72205, USA.

e-mail: shaughnessyjohn@umas.edu

1. Felsher, D. W. & Bishop, J. M. *Mol. Cell* **4**, 199–207 (1999).
2. Jain, M. *et al. Science* **297**, 102–104 (2002).
3. Weinstein, I. B. *Science* **297**, 63–64 (2002).
4. Shaffer, A. L. *et al. Nature* **454**, 226–231 (2008).
5. Klein, U. *et al. Nature Immunol.* **7**, 773–782 (2006).
6. Zhan, F. *et al. Blood* **108**, 2020–2028 (2006).
7. Shaughnessy, J. D. Jr *et al. Blood* **109**, 2276–2284 (2007).
8. Bergsagel, P. L. *et al. Blood* **106**, 296–303 (2005).
9. Zhan, F. *et al. Blood* **99**, 1745–1757 (2002).
10. Zhan, F. *et al. Blood* **101**, 1128–1140 (2003).
11. Potter, M. *Immunol. Rev.* **194**, 177–195 (2003).
12. Matsui, W. *et al. Cancer Res.* **68**, 190–197 (2008).
13. Barlogie, B. *et al. Blood* **103**, 20–32 (2004).
14. Hideshima, T., Mitsiades, C., Tonon, G., Richardson, P. G. & Anderson, K. C. *Nature Rev. Cancer* **7**, 585–598 (2007).
15. www.ncbi.nlm.nih.gov/geo

ASTEROIDS

How to make a flying saucer

William F. Bottke

Wherever we look in the Solar System, small bodies often seem to come in twos. Simulations show how asteroids spun in the Sun can produce such pairings — one of whose members acquires a strangely familiar shape.

When I was a kid, I used to stay up on weekends to watch the late-night monster movies on television. As I recall, a fair number of these films involved little green men (or women) arriving in flying saucers. The latest news from the Solar System's recesses (Walsh *et al.*¹, page 188) is that our martian neighbours might have had a little local inspiration in choosing the shape of their craft. It seems that thermal forces cause some asteroids to spin up, change shape and shed mass, and eventually evolve into binary systems with a flying-saucer-shaped primary body.

How does this saucer production line work? The answer is that it is solar powered: the torques that modify the spin rates and pole directions of small asteroids are largely created by sunlight reflected and re-emitted from an asteroid at thermal infrared wavelengths. This is the Yarkovsky–O'Keefe–Radzievskii–Paddack (YORP) effect — a mouthful coined by the geodynamicist David Rubincam in honour of four pioneers of the field^{2,3}. It can slow down the spinning of an asteroid enough that it falls

into a chaotic (tumbling) rotation state, or alternatively spin it up so much that it starts to cast off material from its surface³.

Walsh *et al.* now report a new addition to YORP's repertoire of outcomes: for kilometre-sized rubble-pile asteroids, their numerical simulations show that the YORP effect can knead the object so that its constituent boulders slide or roll over one another, moving from the poles towards the equator. Over time, this mass movement creates an oblate (slightly squashed) spinning-top-shaped object (Fig. 1, overleaf). This body ends up with bald spots at its poles, where underlying bedrock might be seen if it exists, and a conspicuously fat equatorial belt from which fragments can escape into nearly circular orbits. The high likelihood that these flying saucers have freshly exposed surfaces would make them intriguing targets for our own spacecraft missions.

The authors' simulations further show that fragments ejected from these objects mostly end up accumulating into a single satellite that grows over time. This satellite evolves in



Figure 1 | Science fact. The spinning-up of celestial bodies by various mechanisms to make flying-saucer-shaped objects, identified by Walsh *et al.*¹ in the progenitors of some binary asteroids, has amazing precedents in the literature. This artwork, by Alex Schomburg, comes from the cover of *Amazing Stories*, January 1964 (Ziff-Davis Publishing).

response to the accretion of additional fragments, sunlight-driven, non-gravitational forces, impacts, gravitational kicks from its progenitor and — in the case of ‘near-Earth’ asteroids whose orbits are close to, or even intersect, our own — from close encounters with planets. These mechanisms can also lead to a primary body losing a satellite, but through the YORP effect it readily produces new offspring as long as it has a sufficient reservoir of seed fragments.

A test of Walsh and colleagues’ model¹ is its ability to reproduce the features of the radar-observed binary near-Earth asteroid (NEA) 1999 KY36 (refs 4, 5). The primary body of this binary is a top-shaped asteroid of diameter 1.5 kilometres, with an equatorial belt and a period of rotation — 2.8 hours — that is slightly outside the mass-shedding regime. Its bulk density of just 2 grams per cubic centimetre, estimated by precisely measuring the orbit of the binary and thus its gravitation, indicates that, when compared with meteorite data, the object is 50% porous. This is consistent with its being an ‘unconsolidated’ body made up of loosely packed solid material. The secondary body, with a diameter of 0.5 km, resides on a nearly circular, nearly equatorial orbit 1.5 km from the surface of the primary. Observations of asteroids show that these kinds of diameter ratio, spin rate and orbital distance are common not only among NEA binaries, but also among small-body binaries in the inner ‘main belt’ of asteroids that lies between Mars and Jupiter⁶.

This is not to say that YORP is the only

mechanism creating binary asteroids: they can also arise through impacts, and NEA binaries in particular can originate in the splitting of rubble-pile asteroids into gravitationally bound fragments during a close encounter with a planet^{7,8}. Whereas such ‘tidal’ disruption seems considerably less efficient at making NEA binaries than YORP is, impacts might be a more competitive mechanism. Numerical simulations indicate that large asteroid disruptions produce numerous ‘escaping ejecta binaries’ — fragments launched from the collision site that have similar enough trajectories that they become gravitationally bound together⁸. Such binaries, moving out of the main belt and into the NEA population, would look very different from the flying-saucer binaries.

The armada of flying saucers has outliers elsewhere in the Solar System. Pan and Atlas, two small satellites embedded within the rings of Saturn, also have prominent equatorial ridges, produced by ring particles that struck and piled up on their surfaces⁹. But the mother ship of all of the flying saucers has to be Iapetus. This 740-km-diameter, walnut-shaped moon of Saturn has a protruding waistline, squashed poles, and a discontinuous equatorial ridge that is 18 km high, 200 km wide and 1,600 km long^{10,11}. The oblate shape and prominent ridge of Iapetus are reminiscent of features on the smaller flying saucers, and their formation mechanisms might share similarities as well — produced perhaps by a fast spin early in Iapetus’ history¹¹ and/or collisional accretion of particles within a localized disk around the moon¹².

Flying saucers are not the alpha and omega of asteroid shapes and sizes. Since the groundbreaking discovery of a main-belt asteroid with its own satellite (Ida and its moon Dactyl), ground- and space-based searches using state-of-the-art observational techniques have turned up satellite systems in every population of small bodies that we know of — from main-belt asteroids to near-Earth asteroids to Trojans (asteroids that share Jupiter’s orbit) to the bodies of the Kuiper belt beyond Neptune⁷. By understanding these populations, what they are like, and how they formed, we can glean insights into the evolutionary history of small bodies throughout the Solar System — although I suspect we are fated never to find one actually carrying those little green men. ■

William F. Bottke is in the Department of Space Studies, Southwest Research Institute, 1050 Walnut Street, Boulder, Colorado 80302, USA. e-mail: bottke@boulder.swri.edu

1. Walsh, K. J., Richardson, D. C. & Michel, P. *Nature* **453**, 188–191 (2008).
2. Rubincam, D. P. *Icarus* **148**, 2–11 (2000).
3. Bottke, W. F., Vokrouhlický, D., Rubincam, D. P. & Nesvorný, D. *Annu. Rev. Earth Planet. Sci.* **34**, 157–191 (2006).
4. Ostro, S. J. *et al.* *Science* **314**, 1276–1280 (2006).
5. Scheeres, D. J. *et al.* *Science* **314**, 1280–1283 (2006).
6. Pravec, P. & Harris, A. W. *Icarus* **190**, 250–259 (2007).
7. Richardson, D. C. & Walsh, K. J. *Annu. Rev. Earth Planet. Sci.* **34**, 47–81 (2006).
8. Durda, D. D. *et al.* *Icarus* **167**, 382–396 (2004).
9. Charoz, S., Brahic, A., Thomas, P. C. & Porco, C. C. *Science* **318**, 1622–1624 (2007).
10. Porco, C. C. *et al.* *Science* **307**, 1237–1242 (2005).
11. Castillo-Rogez, J. C. *et al.* *Icarus* **190**, 179–202 (2007).
12. Ip, W.-H. *Geophys. Res. Lett.* **33**, L16203 (2006).

SOLID-STATE CHEMISTRY

New order for lithium bromide

David C. Johnson

It’s all very well predicting which forms of crystal an inorganic solid can adopt, but how can proof be obtained if these structures aren’t thermodynamically stable? The answer is to build them up atom by atom.

Some might say that organic chemists have it easy. When they run a reaction, they can be reasonably confident that the product will be the one they predicted. And if the reaction doesn’t proceed as expected, then at least the products will usually be structurally related to the starting materials. Solid-state chemists have no such luxury — the prediction of the atomic arrangements adopted by inorganic structures is more challenging because of the wide variation in structural motifs and geometries that can be adopted by combinations of inorganic atoms¹. But this looks set to change. Having devised computational methods to predict the atomic configurations of metastable solids², Jansen and colleagues³ now report in *Angewandte Chemie* the preparation of a

metastable form of lithium bromide (LiBr), in a process that avoids the formation of thermodynamically stable arrangements.

Quite apart from the difficulties in predicting the extended (lattice) structures of solid-state compounds, practical challenges mark out this field from that of molecular synthesis. When making a molecule, reactions are typically performed in solution, where high diffusion rates allow the reactant molecules to move around and encounter each other readily. The rate-limiting step in these reactions is typically the breaking of a specific chemical bond (which can be tuned by the proper choice of a catalyst), and a dynamic chemical equilibrium exists between the reactants and products. By contrast, the synthesis of solid-state structures typically involves

a heterogeneous reaction mixture of two or more solids; the diffusion rates of atoms in solids at their reaction temperatures can be billions of times slower than diffusion rates in liquids, so the rate-limiting step is often the interdiffusion of the reactants (which can't be speeded up by catalysts). And because the extended structure of the solid product is typically unrelated to that of the starting materials, the reactions are irreversible.

As if all this weren't difficult enough, the preparation of metastable, extended solid-state structures adds another level of complexity: the nucleation of the new solid requires tens to hundreds of atoms to self-assemble into a nucleus with the desired crystal structure. This must then be followed by the rearrangement of about 10^{23} additional atoms as the product crystals grow. All of this must happen under conditions that prevent the formation of the preferred, thermodynamically stable phase⁴.

Jansen and colleagues³ have found a way to address these issues. The authors had previously used a computational approach to predict that several crystal forms of lithium bromide would be more stable than its thermodynamically preferred structure⁵ — the rock-salt structure (Fig. 1a) — if the system could be forced to have a lower density. They now show³ that one of the predicted structures (known as a wurtzite arrangement; Fig. 1b) can indeed be formed by depositing the elements of lithium bromide atom by atom onto a substrate at low temperatures.

The authors began their search for a new form of lithium bromide by preparing an amorphous solid that has the same elemental composition as the desired product. This isn't a new idea — amorphous phases of compounds have previously been used as precursors to metastable, crystalline forms of those compounds. In amorphous solids, the atoms are uniformly distributed and dispersed, thus avoiding the complications of heterogeneous reaction mixtures and reducing the distances that atoms have to travel by diffusion. In general, the phase that crystallizes from an amorphous precursor is the one that nucleates most easily, and not necessarily the one that is thermodynamically most stable.

The energetic barrier to nucleation is the result of a trade-off between the 'surface energy' required for the new arrangement of atoms to grow in the matrix of the precursor, and the relative volume stability of the phase being formed compared with that of the precursor. Several other factors determine whether or not a particular metastable phase can be isolated, including the dependence of the surface energy on the solid's change in volume when it converts from an amorphous to a crystalline state. The exact phase that nucleates from an amorphous intermediate also depends on the atomic composition of that phase, and the local stresses and strains that develop as the crystal forms⁶.

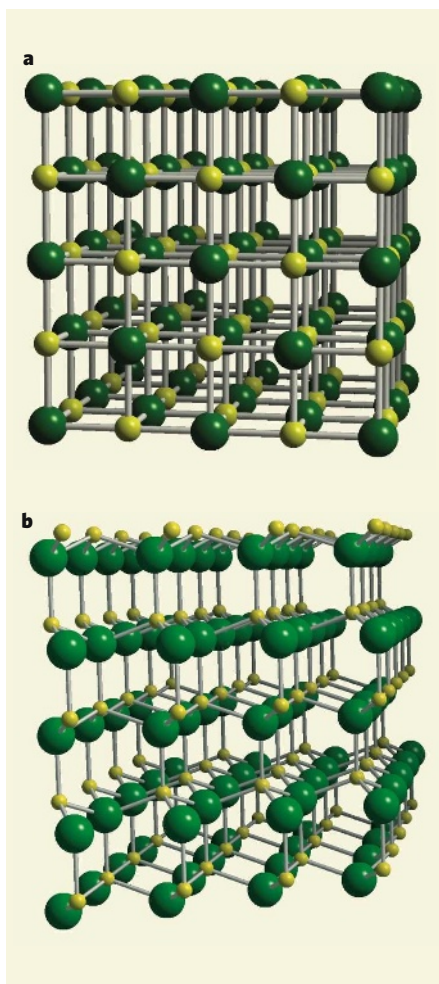


Figure 1 | A fresh look for lithium bromide.

a, The thermodynamically preferred arrangement of atoms in lithium bromide is the 'rock-salt' formation. **b**, Jansen and colleagues³ show that a metastable form of lithium bromide, known as a wurtzite arrangement of atoms, is formed if the elements of lithium bromide are deposited atom by atom on a cold substrate.

Jansen and colleagues³ found that when they warmed their amorphous lithium bromide, the thermodynamically preferred rock-salt structure, rather than the metastable wurtzite phase, nucleated and grew. This was because diffusion of the atoms in the amorphous solid occurred only at temperatures that were also high enough to allow the system to reach equilibrium (at which point no metastable structures will form). The authors therefore explored different substrate temperatures and elemental deposition rates in an attempt to find conditions in which nucleation of the wurtzite phase occurs preferentially. Substrate temperatures control atomic diffusion rates on the surfaces of growing crystals, whereas the vapour pressure of the depositing source determines how long atoms spend at the surface before being covered by freshly deposited material. These parameters influence both the nucleation of structures and the subsequent growth of compounds that have nucleated.

The authors found that the wurtzite phase of

lithium bromide forms uniquely at $-50\text{ }^{\circ}\text{C}$, as long as the vapour pressure of the atoms being deposited on the substrate is low (2.3×10^{-4} millibar, thousands of times lower than atmospheric pressure). They speculate that, under their experimental conditions, the strain that is released when the amorphous lithium bromide changes to a crystalline state favours the formation of the wurtzite structure rather than the denser rock-salt arrangement of atoms. Curiously, higher or lower temperatures or vapour pressures resulted in smaller amounts of the wurtzite phase crystallizing.

This work³ demonstrates beautifully that extended solid-state structures can be predicted and prepared. Nevertheless, practical challenges remain before this strategy can be extended to other molecular systems. Lithium bromide is a relatively simple system, because it contains only two types of atom in equal proportions. But computational predictions of structures become much more difficult as the number of atoms in the basic molecular unit increases, and as the ratios of those atoms vary. Fortunately, developing computationally efficient ways of predicting stable configurations of atoms is a hot topic of current research, with applications in several fields — such methods are crucial for understanding protein folding, for example. The synthesis of metastable systems consisting of more than two types of atom also becomes more challenging, as many more thermodynamically stable arrangements of atoms are available, which must be avoided.

Understanding crystal nucleation remains a crucial scientific challenge: if nucleation could be controlled, many technological benefits would follow. Potential applications include such diverse areas as stabilizing amorphous alloys, controlling particle size and density in nanocomposites, manipulating the average particle size and size distributions in nanocrystalline materials prepared from solution, and the synthesis of new stable and metastable compounds that might have interesting material properties. Jansen and colleagues' findings³ highlight the need to explore and control several experimental conditions, to develop new synthetic approaches, and to devise more detailed mechanisms for the major steps in solid-state reactions. ■

David C. Johnson is in the Department of Chemistry and the Materials Science Institute, University of Oregon, Eugene, Oregon 97403, USA.
e-mail: davej@uoregon.edu

- DiSalvo, F. J. *Science* **247**, 649–655 (1990).
- Schön, J. C. & Jansen, M. *Angew. Chem. Int. Edn* **35**, 1286–1304 (1996).
- Liebold-Ribeiro, Y., Fischer, D. & Jansen, M. *Angew. Chem. Int. Edn* **47**, 4428–4431 (2008).
- Stein, A., Keller, S. W. & Mallouk, T. E. *Science* **259**, 1558–1564 (1993).
- Čančarevič, Ž. P., Schön, J. C. & Jansen, M. *Chem. Asian J.* **3**, 561–572 (2008).
- Brophy, J. H., Rose, R. M. & Wulff, J. *The Structure and Properties of Materials* (Wiley, 1964).

Do female hyaenas choose mates based on tenure?

Arising from: O. P. Höner *et al.* *Nature* **448**, 798–801 (2007)

In their investigation into whether female mate-choice drives male dispersal, Höner *et al.*¹ argue that female spotted hyaenas (*Crocuta crocuta*) prefer mates whose tenure in the social group is less than the females' age, to avoid paternal incest, and suggest that male dispersal reflects this preference. However, we are not persuaded that females choose mates on the basis of tenure because Höner *et al.*¹ overlook the alternative hypothesis that dispersal status itself is important in female mate-choice^{2,3}, such that females prefer immigrants over natal males. Like mate-choice based on tenure, choice based on dispersal status reduces the risk of incest.

Female hyaenas discriminate immigrants from natal males⁴, so if natal males sire offspring in proportion to their abundance, this argues against female choice based on dispersal status. Behaviourally, 11% of Ngorongoro Crater males "started their reproductive career" in their natal groups¹. On the basis of comparable behaviours, 68% of males born in a Masai Mara population also did so⁵. (Note that this *Crocuta* population is the same one from which we derived the data for our Fig. 1.) Nevertheless, adult natal males were less likely to father any cubs than were immigrants² and natal males sired only 2.7% of cubs, although they comprised more than 20% of adult males² ($\chi^2 = 16.04$, degrees of freedom, d.f. = 1, $P < 0.001$). Thus, prior paternity data² indicate that females avoid mating with natal males, preferring immigrants.

Mate-choice in hyaenas often seems, perhaps coincidentally, to reflect both choice based on tenure and choice based on dispersal status. Of 76 litters born in the Masai Mara study population (see above), 51 (67.1%) were sired by males consistent with both choice based on tenure and dispersal status (Fig. 1). Therefore, we focus on the remaining 25 litters, which allow us to distinguish between the two competing hypotheses. Of these 25 litters, 22 (88%) were sired by immigrants present in the clan when the mothers were born, supporting choice based on dispersal status, but conflicting with choice based on tenure. However, only 3 out of 25 litters (12%) were sired by natal males born after the mothers, supporting tenure-based choice while conflicting with choice based on dispersal status. Thus, most of our informative litters support female choice based on dispersal status, and conflict with choice based on tenure. Although most (89.6%)

Ngorongoro Crater litters seem to support tenure-based mate-choice¹, Höner *et al.*¹ present no paternity data comparing reproductive success between immigrants and natal males. Without such data, we cannot know whether Ngorongoro Crater females prefer immigrants as mates.

Furthermore, Fig. 1 and related analyses of Höner *et al.*¹ may derive from female preference for immigrants², coupled with age differences between natal females and immigrant males. If litters in Fig. 1 of Höner *et al.*¹ were sired primarily by immigrants, as indicated by our data, and Ngorongoro Crater males immigrate when about 4 years old, as elsewhere^{6,7}, then their Fig. 1 and ours both compare females of age X years with males aged $(X + 4)$ years. Males survive no longer than females⁸, so both figures may reflect a dearth of old males, not tenure-based mate-choice.

We do not suggest that females in this or other species choose mates based solely on male dispersal status. However, male-biased dispersal is common among mammals^{9,10}, so dispersal-based mate-choice could occur in many species. By contrast, tenure-based choice is possible only in social systems like that of *Crocuta*, so it cannot explain widespread sex differences in dispersal behaviour. In addition, because spotted hyaenas discriminate maternal kin, paternal kin, and sires from non-kin^{11,12}, females might choose mates based on several criteria. However, the role of male tenure in female mate-choice¹ remains unclear.

METHODS

We quantified female mate-choice in a large Kenyan social group by classifying hyaenas as adults at 24 months of age¹, considering immigrants as residents 3 months after their arrival¹, and assigning kinship by paternity analysis². We assigned paternity to 76 out of 110 litters born in 1989–2000. Five litters were sired by multiple males. Such litters occur near Ngorongoro Crater⁶, but it is unclear how Höner *et al.*¹ treated such litters; for these, we plot the sires' average tenure.

Russell C. Van Horn^{1,2}, Heather E. Watts² & Kay E. Holekamp²

¹Zoological Society of San Diego, Conservation and Research for Endangered Species, P.O. Box 120551, San Diego, California 92112-0551, USA.

e-mail: rvanhorn@sandiegozoo.org

²Department of Zoology, Michigan State University, East Lansing, Michigan 48824-1115, USA.

Received 29 October 2007; accepted 17 April 2008.

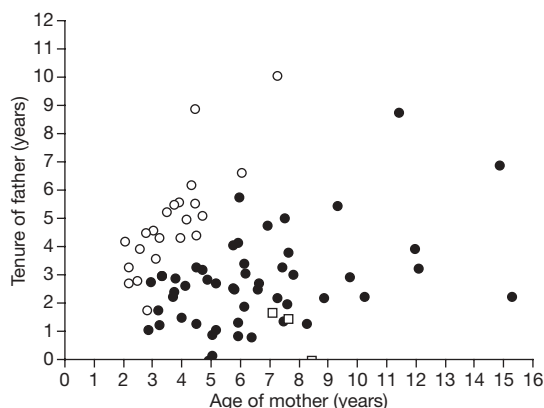


Figure 1 | The relationship between age of the mother on the date of conception, the tenure of the sire, and the sire's dispersal status. ($N = 76$ litters.) Filled circles, litters sired by immigrants who joined the clan after the mother was born, supporting mate-choice based on both dispersal status and tenure. Open circles, litters sired by immigrants who were present in the clan when the mother was born, supporting choice based on dispersal status but not tenure. Open squares, litters with a natal male sire younger than the mother, supporting choice based on tenure but not dispersal status.

- Höner, O. P. *et al.* Female mate-choice drives the evolution of male-biased dispersal in a social mammal. *Nature* **448**, 798–801 (2007).
- Engh, A. L. *et al.* Reproductive skew among males in a female-dominated mammalian society. *Behav. Ecol.* **13**, 193–200 (2002).
- Lehmann, L. & Perrin, N. Inbreeding avoidance through kin recognition: choosy females boost male dispersal. *Am. Nat.* **162**, 638–652 (2003).
- Szykman, M., Van Horn, R. C., Engh, A. L., Boydston, E. E. & Holekamp, K. E. Courtship and mating in free-living spotted hyenas. *Behaviour* **144**, 815–846 (2007).
- Holekamp, K. E. & Smale, L. Dispersal status influences hormones and behavior in the male spotted hyena. *Horm. Behav.* **33**, 205–216 (1998).
- East, M. L. & Hofer, H. Male spotted hyenas (*Crocuta crocuta*) queue for status in social groups dominated by females. *Behav. Ecol.* **12**, 558–568 (2001).
- Van Horn, R. C., McElhinney, T. L. & Holekamp, K. E. Age estimation and dispersal in the spotted hyena (*Crocuta crocuta*). *J. Mamm.* **84**, 1019–1030 (2003).
- Boydston, E. E., Kapheim, K. M., Van Horn, R. C., Smale, L. & Holekamp, K. E. Sexually dimorphic patterns of space use throughout ontogeny in the spotted hyena (*Crocuta crocuta*). *J. Zool.* **267**, 271–281 (2005).
- Greenwood, P. J. Mating systems, philopatry and dispersal in birds and mammals. *Anim. Behav.* **28**, 1140–1162 (1980).
- Lawson Handley, L. J. & Perrin, N. Advances in our understanding of mammalian sex-biased dispersal. *Mol. Ecol.* **16**, 1559–1568 (2007).
- Wahaj, S. A. *et al.* Kin discrimination in the spotted hyena (*Crocuta crocuta*): nepotism among siblings. *Behav. Ecol. Sociobiol.* **56**, 237–247 (2004).
- Van Horn, R. C., Wahaj, S. A. & Holekamp, K. E. Role-reversed nepotism among cubs and sires in the spotted hyena (*Crocuta crocuta*). *Ethology* **110**, 413–426 (2004).

doi:10.1038/nature07122

Höner et al. reply

Replying to: R. C. Van Horn, H. E. Watts & K. E. Holekamp *Nature* 454, doi:10.1038/nature07122 (2008)

We demonstrated that female mate-choice, rather than male inbreeding avoidance, resources or male–male competition, drives male-biased dispersal in spotted hyaenas (*Crocuta crocuta*)¹. We further showed that females use two simple rules based on male tenure to choose their sires, and that males gain substantial fitness benefits by responding to these preferences¹. Van Horn *et al.*² accept our principal conclusions but claim that females base mate-choice on dispersal status rather than tenure. Their argument overlooks the fact that at least one of the two female mate-choice rules cannot be explained by dispersal status, and it is based on statistical analyses and a selection of males that we consider inappropriate for tests of female mate preference.

Young female spotted hyaenas prefer to breed with short-tenured rather than long-tenured males^{1,3}. This preference cannot be explained by a rule based on dispersal status but must involve tenure-based rules such as those suggested by us¹.

In spotted hyaenas, many natal males do not display sexual interest in females before they disperse^{1–4}. Because females are unlikely to consider such males as potential mates, they should be excluded from tests of female mate preferences^{1,3,5,6}, as is common practice among primatologists^{5,6}. Any appropriate test of female mate preference then considers the number of reproductively active natal males and immigrant males at the conception of each litter and averages the results over all litters per female to yield one data point per female^{1,7}. The comparisons by Van Horn *et al.*² of the percentage success of natal males and immigrant males in the pools of all natal males and all immigrant males present during their study do not recognize these aspects.

First, their tests of preference include natal males without considering their reproductive activity. Their² quote of 68%⁸ of natal males in their *Crocuta* study group (in contrast to 11% in our study population¹) showing reproductive behaviour to females is puzzling given that testosterone concentrations of natal males resemble testosterone concentrations of juvenile non-reproductive males⁹. In the paternity analysis of Van Horn *et al.*² (citing ref. 10), natal males are assumed not to be candidates and therefore presumptively excluded if they are a relative (at $r \geq 0.125$) of the female, thereby reducing the chance of natal males being identified as sires. Thus, the tests of preference by Van Horn *et al.*² and a previous paternity study on the same clan¹⁰ compare the reproductive success of unrelated natal males with a pool comprising reproductively active and inactive, sexually immature, related and unrelated natal males, thereby reducing the percentage of successful natal males and favouring immigrant males^{6,7}.

Second, Van Horn *et al.*² did not take into account the fact that male reproductive success strongly increases with tenure^{3,10}. With their method, reproductive success becomes skewed towards long-tenured immigrant males¹⁰ owing to significant differences in mean

tenure between immigrant males and natal males^{3,10}. This demonstrates the importance of tenure, not dispersal status.

Third, for their 22 litters apparently not conforming to the tenure-based rule, Van Horn *et al.*² do not provide any evidence or test that compares the availability of immigrant males and reproductively active natal males on the dates of conception, nor the availability of males born/immigrated before and after the female's birth.

Van Horn *et al.*² state that our result that females avoid males present when the females are born¹ may reflect a lack of such male candidates. This is incorrect because our test considered the proportion of candidate males that initiated their reproductive career before and after the females were born¹. And if we consider only those litters where females can choose between both males that initiated their tenure before and after the female's birth, they still strongly avoid males present when they were born (Wilcoxon signed-rank test, $n = 54$ females, $P = 0.0001$).

Do females prefer immigrant males in our eight study groups? No: the reproductive success of natal males during their reproductive tenure¹ did not differ from the mean reproductive success of their immediate immigrant male predecessor and successor during the same duration of tenure (Wilcoxon signed-rank test, $n = 12$, exact $P = 0.148$, power = 0.35).

Oliver P. Höner¹, Bettina Wachter¹, Marion L. East¹, W. Jürgen Streich¹, Kerstin Wilhelm¹, Terry Burke² & Heribert Hofer¹

¹Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Strasse 17, D-10315 Berlin, Germany.

e-mail: hoener@izw-berlin.de

²Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield S10 2TN, UK.

- Höner, O. P. *et al.* Female mate-choice drives the evolution of male-biased dispersal in a social mammal. *Nature* **448**, 798–801 (2007).
- Van Horn, R. C. *et al.* Do female hyaenas choose mates based on tenure? *Nature* **454**, doi:10.1038/nature07122 (2008).
- East, M. L. *et al.* Sexual conflicts in spotted hyenas: male and female mating tactics and their reproductive outcome with respect to age, social status and tenure. *Proc. R. Soc. Lond. B* **270**, 1247–1254 (2003).
- East, M. L. & Hofer, H. Male spotted hyenas (*Crocuta crocuta*) queue for status in social groups dominated by females. *Behav. Ecol.* **12**, 558–568 (2001).
- Bercovitch, F. B. Male rank and reproductive activity in savanna baboons. *Int. J. Primatol.* **7**, 533–550 (1986).
- Alberts, S. C., Buchan, J. C. & Altmann, J. Sexual selection in wild baboons: from mating opportunities to paternity success. *Anim. Behav.* **72**, 1177–1196 (2006).
- Manly, B. F. J., McDonald, L. L. & Thomas, D. L. *Resource Selection by Animals. Statistical Design and Analysis for Field Studies* (Chapman & Hall, London, 1993).
- Holekamp, K. E. & Smale, L. Dispersal status influences hormones and behaviour in the male spotted hyena. *Horm. Behav.* **33**, 205–216 (1998).
- Holekamp, K. E. & Sisk, C. L. Effects of dispersal status on pituitary and gonadal function in the male spotted hyena. *Horm. Behav.* **44**, 385–394 (2003).
- Engh, A. L. *et al.* Reproductive skew among males in a female-dominated mammalian society. *Behav. Ecol.* **13**, 193–200 (2002).

doi:10.1038/nature07123

Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor

Jeffrey E. Lee¹, Marnie L. Fusco¹, Ann J. Hessel¹, Wendelien B. Oswald¹, Dennis R. Burton¹ & Erica Ollmann Saphire¹

Ebola virus (EBOV) entry requires the surface glycoprotein (GP) to initiate attachment and fusion of viral and host membranes. Here we report the crystal structure of EBOV GP in its trimeric, pre-fusion conformation (GP1+GP2) bound to a neutralizing antibody, KZ52, derived from a human survivor of the 1995 Kikwit outbreak. Three GP1 viral attachment subunits assemble to form a chalice, cradled by the GP2 fusion subunits, while a novel glycan cap and projected mucin-like domain restrict access to the conserved receptor-binding site sequestered in the chalice bowl. The glycocalyx surrounding GP is likely central to immune evasion and may explain why survivors have insignificant neutralizing antibody titres. KZ52 recognizes a protein epitope at the chalice base where it clamps several regions of the pre-fusion GP2 to the amino terminus of GP1. This structure provides a template for unravelling the mechanism of EBOV GP-mediated fusion and for future immunotherapeutic development.

The EBOV is an enveloped, non-segmented, negative-strand RNA virus, which with Marburg virus makes up the *Filoviridae* family. The virus causes severe haemorrhagic fever associated with 50–90% human mortality¹. Four species of the virus (Zaire, Sudan, Côte d'Ivoire and Reston ebolavirus) have thus far been identified, with the Zaire species typically associated with the highest human lethality². A fifth species was confirmed in a 2007 outbreak in Bundibugyo, Uganda^{3,4}. Infection with EBOV results in uncontrolled viral replication and multiple organ failure, with death occurring 6–9 days after onset of symptoms⁵. Fatal cases are associated with high viraemia and defective immune responses, whereas survival is associated with early and vigorous humoral and cellular immune responses^{6–9}. Although preliminary vaccine trials in primates have been highly successful^{10–13}, no vaccines, specific immunotherapeutics or post-exposure treatments are currently approved for human use. Since 1994, EBOV outbreaks have increased more than fourfold, thus necessitating the urgent development of vaccines and therapeutics for use in the event of an intentional, accidental or natural EBOV release.

The EBOV genome contains seven genes, which direct the synthesis of eight proteins. Transcriptional editing of the fourth gene (GP) results in expression of a 676-residue transmembrane-linked glycoprotein termed GP, as well as a 364-residue secreted glycoprotein termed sGP^{14,15}. EBOV GP is the main target for the design of vaccines and entry inhibitors. GP is post-translationally cleaved by furin¹⁶ to yield disulphide-linked GP1 and GP2 subunits¹⁷. GP1 effects attachment to host cells, whereas GP2 mediates fusion of viral and host membranes^{16,18–20}. EBOV is thought to enter host cells by receptor-mediated endocytosis through clathrin-coated pits and caveolae²¹, followed by actin- and microtubule-dependent transport to the endosome²¹, where GP is further processed by endosomal cathepsins^{22–24}. Essential cellular receptor(s) have not yet been identified, but DC-SIGN/L-SIGN²⁵, hMGL²⁶, β -integrins²⁷, folate receptor- α ²⁸ and Tyro3 family receptors²⁹ have all been implicated as cellular factors in entry. Here we report the crystal structure of

EBOV GP, at 3.4-Å resolution, in its trimeric, pre-fusion conformation in complex with neutralizing antibody Fab KZ52.

EBOV GP is in a pre-fusion conformation

In an effort to increase sample homogeneity and to promote crystal contacts, we excised the mucin-like and transmembrane domains (GP_{33–632} Δ muc), mutated two N-linked glycosylation sites (T42V/T230V) and complexed the GP with Fab KZ52, which recognizes a conformational epitope. The resulting GP construct is fully capable of mediating virus entry and exhibits similar antibody neutralization profiles as wild type, when expressed with a transmembrane domain on vesicular stomatitis virus pseudovirions (Supplementary Methods and Fig. 1).

The EBOV GP trimer contains three non-covalently attached monomers (A, B and C) (Supplementary Fig. 2), which together adopt a chalice-like shape with overall dimensions of about 95 Å × 95 Å × 70 Å (Fig. 1). Each monomer consists of two disulphide-linked subunits, GP1 and GP2. GP trimerization is mediated through multiple GP1–GP2 and GP2–GP2 contacts and occludes an approximate 2,900 Å² surface area on each subunit. No major contacts are observed between neighbouring GP1 molecules. Three GP1 ectodomains together form a bowl-like structure encircled by helices of the three GP2 subunits (Fig. 1c).

EBOV GP is thought to assemble as a metastable trimer of heterodimers on the viral surface. Upon a yet unidentified trigger, GP undergoes an irreversible conformational change to a lower energy state, thereby merging the viral and host cell membranes. The low energy or post-fusion conformation is characterized by a six-helix bundle, in which the N terminus and fusion loop of GP2 are juxtaposed to its carboxy terminus in a hairpin conformation^{30,31}. Our EBOV GP structure likely corresponds to the metastable, pre-fusion conformation of EBOV GP, as GP2 forms a trimer intimately associated with GP1, rather than the post-fusion six-helix bundle conformation. In addition, crystals were grown at a neutral pH and the EBOV GP presented here contains no multimerization motifs

¹Department of Immunology and Microbial Science, The Scripps Research Institute, 10550 North Torrey Pines Road, Mail Drop IMM-2, La Jolla, California 92037, USA.

or recombinant constraints on quaternary structure that could artificially force the structure into a non-prefusion state.

EBOV GP1

GP1 is responsible for cell-surface attachment, which is probably mediated by a region including residues 54–201 (ref. 32). It is composed of a single domain (approximately $65 \text{ \AA} \times 30 \text{ \AA} \times 30 \text{ \AA}$), arranged in the topology shown in Fig. 2a, and can be further subdivided into the (I) base, (II) head and (III) glycan cap regions (Fig. 2a and Supplementary Fig. 3). The base (I) subdomain is composed of two sets of β -sheets, forming a semicircular surface that clamps the internal fusion loop and a helix of GP2 through hydrophobic interactions (Fig. 2b). Moreover, this subdomain contains Cys 53, which is proposed to form an intermolecular disulphide bridge to Cys 609 of the GP2 subunit¹⁷. Cys 53 resides near GP2 in the $\beta 2$ – $\beta 3$ loop at the viral membrane-proximal end of the base subdomain (Fig. 2a, b). Our EBOV GP contains an intact GP1–GP2 disulphide bridge, based on reducing and non-reducing SDS–polyacrylamide gel electrophoresis. However, the region containing the counterpart GP2 cysteine is disordered, which may reflect functionally important mobility in the region. The head (II) is located between the base and glycan cap regions towards the host membrane surface. Two intramolecular disulphide bonds stabilize the head subdomain and confirm the biochemically determined disulphide bridge assignments¹⁷. Cys 108–Cys 135 connects a surface-exposed loop ($\beta 8$ – $\beta 9$ loop) to strand $\beta 7$, whereas Cys 121–Cys 147 bridges the $\beta 8$ – $\beta 9$ and $\beta 9$ – $\beta 10$ loops (Fig. 2a). The glycan cap (III) contains four predicted N-linked glycans (at N228, N238, N257 and N268) in an α/β -dome over the GP1 head subdomain (Figs 1b and 2a). This subdomain does not form any monomer–monomer contacts and is fully exposed on the upper and outer surface of the chalice. The central β -sheets from the head and glycan cap together form a fairly flat surface and, in the context of the GP trimer, form the three inner sides of the chalice bowl.

EBOV GP2

GP2 is responsible for fusion of viral and host cell membranes and contains the internal fusion loop and the heptad repeat regions, HR1 and HR2. Many viral glycoproteins have fusion peptides, located at the N terminus of their fusion subunit, which are released upon cleavage of the precursor glycoprotein. In contrast, class II and class III fusion proteins, as well as class I glycoproteins from Ebola, Marburg, Lassa and avian sarcoma leukosis viruses, contain internal fusion loops lacking a free N terminus. The crystal structure reveals that the EBOV GP internal fusion loop, which encompasses residues 511–556, uses an antiparallel β -stranded scaffold to display a partly helical hydrophobic fusion peptide (L529, W531, I532, P533, Y534 and F535; Fig. 2c). The side chains of these hydrophobic residues pack into a region on the GP1 head of a neighbouring subunit in the trimer, reminiscent of the fusion peptide packing in the pre-fusion parainfluenza virus 5 F structure³³. A disulphide bond between Cys 511 at the base of $\beta 19$ and Cys 556 in the HR1 helix covalently links the antiparallel β -sheet. This disulphide bond between the internal fusion loop and HR1 is conserved among all filoviruses, and is analogous to a pair of critical cysteines flanking the internal fusion loop in avian sarcoma leukosis virus^{34,35}. Interestingly, the EBOV internal fusion loop has features more similar to those observed in class II and III viral glycoproteins (in particular to flaviviruses) than those previously observed for class I glycoproteins (Supplementary Fig. 4). It thus appears that regardless of viral protein class, internal fusion loops share a common architecture for their fusion function.

EBOV GP2 contains two heptad repeat regions (HR1 and HR2), connected by a 25-residue linker containing a CX₆CC motif and the internal fusion loop. The crystal structures of post-fusion GP2 fragments^{30,31} have revealed that the two heptad repeat regions form antiparallel α -helices and that a CX₆CC motif forms an intrasubunit disulphide bond between Cys 601 and Cys 608 (Supplementary Fig.

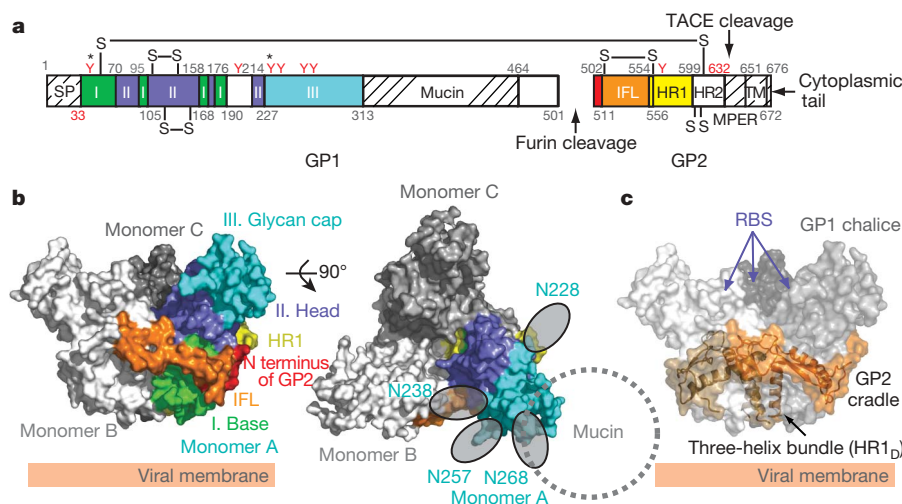


Figure 1 | Structure of Zaire EBOV GP. **a**, Domain schematic of GP. Domains observed in the crystal structure are coloured and numbered according to the description in the text. White and hash-marked regions designate crystallographically disordered and construct-deleted regions, respectively. SP, signal peptide; I, GP1 base; II, GP1 head; III, GP1 glycan cap; mucin, mucin-like domain; IFL, internal fusion loop; HR1, heptad repeat 1; HR2, heptad repeat 2; MPER, membrane-proximal external region; TM, transmembrane domain. Red Y-shaped symbols designate the predicted N-linked glycosylation sites; those sites marked with an asterisk were mutated. The final model includes EBOV GP residues 33–189, 214–278, 299–310 and 502–599. No electron density is observed for residues 190–213, 311–312, 464–501 and 600–632. Weak or discontinuous electron density is seen in the loop containing the GP1–GP2 disulphide bridge (residues 49–56) and the outer regions of the GP1 glycan cap (residues

268–278 and 299–310); these regions are modelled as poly-alanine fragments. **b**, Molecular surface of the GP trimer viewed on its side (left) and top (right), as viewed down the threefold axis. Monomer A is coloured according to the scheme in **a**, and monomers B and C of the trimer are shown in white and grey, respectively. Predicted N-linked oligosaccharides (N228, N238, N257 and N268) belonging to the glycan cap of monomer A are shown as grey ovals. The location of the N-linked glycan at N268 is tentative, as the sequence assignment in this region is ambiguous. **c**, Molecular surface of the EBOV GP chalice and cradle. The three GP1 subunits that form the chalice are shown in various shades of grey, and GP2 subunits forming the cradle are shown as ribbons, in various shades of orange, underneath the transparent molecular surface. The putative RBSs are recessed in the inner bowl of the GP trimer.

5). In the pre-fusion EBOV GP, HR2 and the CX₆CC motif are disordered. In contrast, the HR1 region is well ordered and can be divided into four segments: HR1_A, HR1_B, HR1_C and HR1_D (Fig. 2c), which together assemble the cradle encircling GP1. Similarly, heptad repeat regions in influenza and parainfluenza viruses also contain multiple segments in their pre-fusion helices that substantially rearrange in their post-fusion conformations^{33,36,37}.

The first two segments, HR1_A and HR1_B (residues 554–575), together form an α -helix with an approximate 40° kink at T565, which delineates HR1_A from HR1_B. Interestingly, the bend between HR1_A and HR1_B contains an unusual 3–4–4–3 stutter, which may act as a conformational switch³¹, rather than the typical 3–4 periodicity of heptad repeats (Supplementary Fig. 6). A similar stutter has also been noted in parainfluenza virus 5 F³³. The EBOV HR1_C (residues 576–582) forms an extended coil linking HR1_B to the 14-residue α -helix of HR1_D (residues 583–598). HR1_D forms an amphipathic helix, and the hydrophobic faces of each HR1_D join to form a three-helix bundle at the trimer interface. Although the breakpoint maps directly to a chloride-ion binding site in the post-fusion conformation of GP2^{30,31} and at least two other viruses^{38,39}, no chloride ion is observed here as HR1 and HR2 do not come together to form the

six-helix bundle. Instead, the pre-fusion GP2 adopts a novel conformation, intimately curled around GP1 (Fig. 1c).

EBOV GP–KZ52 interface

KZ52 is an antibody isolated from a human survivor of the 1995 outbreak in Kikwit, Democratic Republic of the Congo (formerly Zaire)⁴⁰. This antibody neutralizes Zaire ebolavirus *in vitro*⁴⁰ and offers protection from lethal EBOV challenge in rodent models⁴¹, but has minimal effects on viral pathogenicity in non-human primates⁴². KZ52 is directed towards a vulnerable, non-glycosylated epitope at the base of the GP chalice, where it engages three discontinuous segments of EBOV GP: residues 42–43 at the N terminus of GP1, and 505–514 and 549–556 at the N terminus of GP2 (Fig. 3 and Supplementary Fig. 7). Although most of the GP surface buried by KZ52 belongs to GP2, the presence of both GP1 and GP2 is critical for KZ52 recognition⁴³. It is likely that GP1 is required to maintain the proper pre-fusion conformation of GP2 for KZ52 binding. Indeed, KZ52 is the only antibody known to bridge both attachment (GP1) and fusion (GP2) subunits of any viral glycoprotein. Given that KZ52 requires a conformational epitope seen only in the GP2 pre-fusion conformation and that the KZ52 epitope is distant from the putative receptor binding site (RBS), KZ52 likely neutralizes by preventing rearrangement of the GP2 HR1_A/HR1_B segments and blocking host membrane insertion of the internal fusion loop. Alternatively, IgG KZ52 may sterically hinder access to the RBS or to a separate binding site of another cellular factor, especially if multiple attachment events are required for entry.

The KZ52 epitope of GP is convex and does not have a high shape complementarity to the antibody (surface complementarity index (Sc) of 0.63), although about 1,600 Å² of each GP monomer are occluded upon KZ52 binding. The antibody contacts a total of 15 GP residues by van der Waals interactions and 8 direct hydrogen bonds (Supplementary Fig. 7). Ten out of 15 residues in the structurally defined KZ52 epitope are unique to Zaire ebolavirus (Supplementary Fig. 6), thus explaining the specificity of KZ52.

EBOV GP glycosylation

We generated a fully glycosylated molecular model of EBOV GP to illustrate the native GP trimer as it exists on the viral surface (Fig. 4). Most N-linked glycosylation sites are concentrated in the mucin-like domain and glycan cap of GP1. Given that the mucin-like domain is about 75 kDa in mass (protein and oligosaccharide), the volume of this domain is predicted to be similar to each GP monomer observed here. The crystal structure suggests that the mucin-like domain is

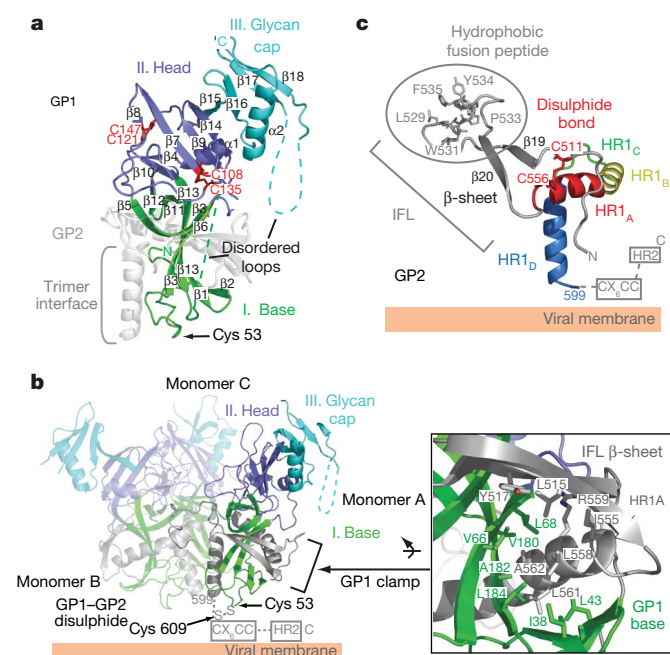


Figure 2 | EBOV GP1 and GP2. **a**, Ribbon diagram of the GP1 subunit. The base (I) subdomain (green) contains four discontinuous sections (residues 33–69, 95–104, 158–167 and 176–189), which form two mixed β -sheets with strands β 3 and β 13 shared between the two β -sheets. The head (II) subdomain (purple) is composed of residues from four discontinuous segments, 70–94, 105–157, 168–175 and 214–226, and forms a four-stranded, mixed β -sheet supported by an α -helix and a smaller, two-stranded, antiparallel β -sheet. The glycan cap (III) region (cyan) is composed of a continuous polypeptide chain, residues 227–310, and forms an α -helix packed against a four-stranded, mixed β -sheet. Intramolecular disulphide bridges between Cys 108–Cys 135 and Cys 121–Cys 147 are coloured red. **b**, Ribbon diagram of the EBOV GP trimer. Each GP1 subdomain is coloured according to panel **a**, and the three GP2 subunits are shown in grey. The GP1 base subdomain forms a clamp onto the GP2 internal fusion loop and HR1_A helix through interactions with hydrophobic residues (inset). The GP1–GP2 disulphide bridge (Cys 53–Cys 609), CX₆CC motif and HR2 region are disordered in the structure and are marked and indicated by dashed lines. **c**, The pre-fusion conformation of EBOV GP2. The EBOV GP2 internal fusion loop contains a disulphide bond at its base (red), antiparallel β strand (grey) and a hydrophobic fusion peptide. The EBOV GP2 HR1 region is segmented into four parts (HR1_A–HR1_D). Note that the HR2 and the CX₆CC motif are disordered in the pre-fusion conformation.

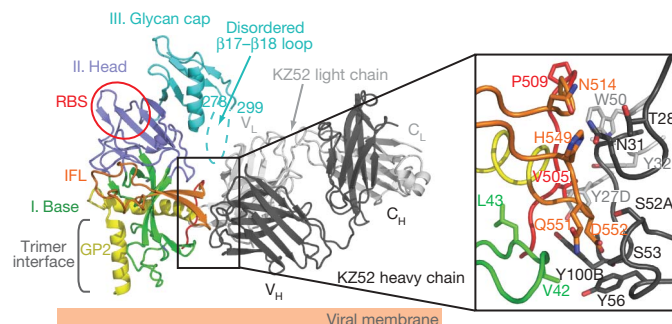


Figure 3 | EBOV GP–Fab KZ52 interactions. KZ52 recognizes residues 505–514 (red) and 549–556 (orange) of GP2 and residues 42–43 (green) of GP1. The β 17– β 18 loop (residues 279–298) is disordered, but may interact with KZ52. One EBOV GP monomer is coloured and labelled according to Fig. 1b, and Fab heavy and light chains are shown in black and light grey, respectively. Side-chain interactions at the GP–KZ52 interface are magnified in the inset box. The putative RBS is outlined in a red circle on the EBOV GP ribbon diagram. Note that only selected residues from GP and KZ52 are shown for clarity, and residue 42 is a threonine in the wild-type Zaire ebolavirus sequence.

linked to the side of each monomer and may further build up the walls of the chalice, forming a deeper bowl (Fig. 4). Although a mixture of complex, oligomannose and hybrid-type glycans are found on intact, mucin-containing GP1⁴⁴, those glycans outside the mucin-like domain are likely to be complex in nature: the mucin-deleted GP used for crystallization is sensitive to PNGaseF, but not to EndoH treatment (Supplementary Fig. 8). Modelling of complex-type oligosaccharides on the EBOV GP indicates that most of the GP trimer is cloaked by a thick layer of oligosaccharide, even without the mucin-like domain (Fig. 4). The approximate 19 additional oligosaccharides on the full-length GP (17 on the mucin-like domain and 2 more on GP1, disordered here) further conceal the sides and top of the chalice. The KZ52 binding site and, presumably, the flexible regions of HR2 and the membrane-proximal external region remain exposed and perhaps vulnerable to binding of antibodies and inhibitors.

The development of neutralizing antibodies is limited in natural EBOV infection. Many survivors have low or insignificant titres^{1,7}, and those antibodies that are elicited preferentially recognize a secreted version of the viral glycoprotein that features an alternative quaternary structure and lacks the mucin-like domain⁴³. The glyco-calyx surrounding EBOV GP likely forms a shield that protects it from humoral immune responses and/or confers stability inside or outside a host. The mucin-like domain and glycan cap sit together as an external domain to the viral attachment and fusion subunits, reminiscent of the glycan shields of HIV-1 gp120 (refs 45 and 46) and Epstein-Barr virus gp350 (ref. 47), perhaps pointing to a common theme for immune evasion. Alignment of filoviral sequences indicate that regions involved in immune evasion have a low degree of sequence conservation (that is, GP1 glycan cap (~5%) and mucin-like domain (0%)), but the N-glycosylation sites in the glycan cap are mostly conserved among all EBOV subtypes (Supplementary Fig. 6), indicating the functional importance of these post-translational modifications.

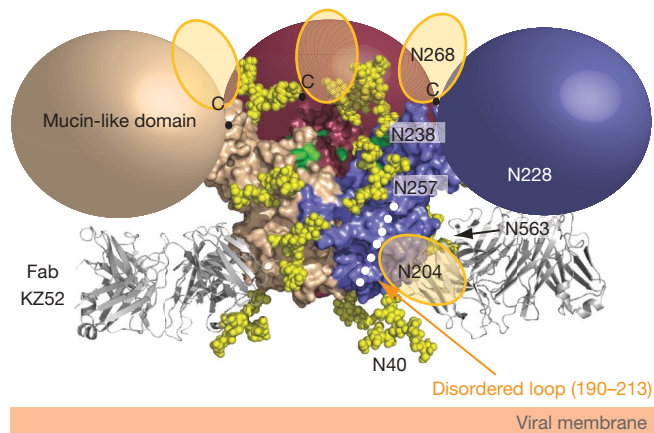


Figure 4 | Model of the fully glycosylated GP. N-linked bi-antennary complex-type glycans ($\text{Gal}_2\text{Man}_3\text{GlcNAc}_4$) were modelled onto the GP1 glycan-cap subdomain. Oligosaccharides are shown as yellow space-filling spheres and for clarity; only those glycans belonging to the purple monomer are labelled. Note that the glycans on N228 and N563 reside on the back of the purple monomer and are partly obscured. The glycans at N204 and N268 are found in regions that are poorly ordered in the structure and as a result: their tentative locations are shown as orange ovals. The C termini of the last ordered residues of GP1, to which mucin-like domains are linked, are marked with 'C' (top of the chalice), and coloured spheres (beige, pink and purple) outline the predicted positions of the mucin-like domains attached in each of these regions. Surface residues previously identified to be critical for viral entry, recessed in the chalice bowl and RBS, are coloured green. Fab KZ52 (grey) recognizes a non-glycosylated, predominantly GP2-containing epitope at the base of the chalice.

Sites of receptor binding and cathepsin cleavage

Although a definitive receptor for EBOV remains to be identified, previous studies^{32,48,49} have determined that residues 54–201, which map to the base and head subdomains of GP1, form a putative RBS for attachment to host cells. Additional studies have identified at least 19 GP1 residues, assigned into four groups based on the location in the structure, that are critical for viral entry^{48–50} (Fig. 5). Many of these residues are apolar or aromatic and are involved in maintaining the structural integrity of GP1 for receptor binding or fusion. However, six residues (K114, K115, K140, G143, P146 and C147) cluster within an approximate $20 \text{ \AA} \times 15 \text{ \AA}$ surface in the inner bowl of the chalice and may thus represent important receptor contact sites. All residues in the putative RBS are highly conserved among EBOV species (Supplementary Fig. 6).

Importantly, this putative RBS is recessed beneath the glycan cap and perhaps further masked by the mucin-like domain (Fig. 4), suggesting that additional conformational change or removal of the mucin-like domain could reveal additional surfaces required for receptor or cofactor binding. It has been demonstrated that endosomal proteolysis of EBOV GP by cathepsin L and/or B removes the mucin-like domain to produce a stable GP1 intermediate of about 18 kDa which has enhanced viral binding and infectivity^{22–24}. The precise site of cathepsin cleavage is unknown and the role of cathepsins in natural infection is as yet unclear. However, formation of an approximate 18 kDa GP1 fragment implies that cathepsin may cleave near the GP1 $\beta 13$ – $\beta 14$ loop (residues 190–213). Indeed, this loop is unresolved in the pre-fusion structure, suggesting enhanced mobility and accessibility to enzymatic cleavage. Cleavage within this loop

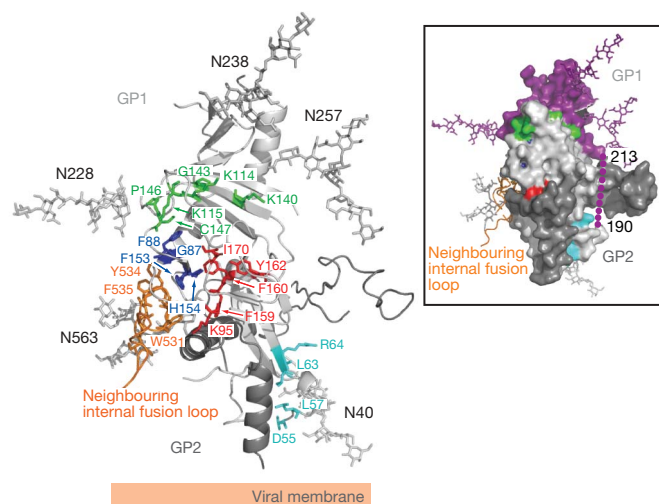


Figure 5 | Sites of receptor binding and cathepsin cleavage. Residues coloured in cyan, green, royal blue and red were previously identified by mutagenesis to be important for viral entry^{48–50}. Residues coloured in cyan (D55, L57, L63 and R64) reside at the base of the chalice, near the GP1–GP2 disulphide bond and HR1_D, and are likely important for fusion-mediated conformational changes rather than receptor binding. Residues coloured in red (F159, F160, Y162 and I170) are primarily buried hydrophobic amino acids that help to maintain the structural stability of GP1. Residues coloured in royal blue (G87, F88, F153 and H154) are in proximity to the putative RBS and pack against the hydrophobic residues from a neighbouring internal fusion loop (coloured orange). Mutations to these residues may affect viral entry by altering the structural integrity of the RBS and/or by affecting packing of the fusion loop. Residues coloured in green (K114, K115, K140, G143, P146 and C147) reside on or near the GP surface and may contribute to receptor binding. A molecular surface representation of a GP monomer, coloured and oriented according to the ribbon diagram, is presented in the inset. Indicated in this view: cleavage at a site on the loop (purple dots) between residues 190 and 213 would remove the entire glycan cap (purple) and the mucin-like domain (not shown), leaving GP2 and an approximate 18 kDa fragment of GP1.

would remove the entire mucin-like domain and glycan cap region (Fig. 5). As a result, $\beta 7$ – $\beta 9$ strands and their associated loops would become exposed. These regions of GP are in proximity to the previously identified residues critical for viral entry. The fold, location and physicochemical properties of this site should now provide new leads in the search for the elusive filoviral receptor(s).

A summary of the EBOV mechanism of infection, including the events of cathepsin cleavage and conformational changes to GP2 during fusion, is presented in Supplementary Information and Supplementary Fig. 9.

METHODS SUMMARY

T42V/T230V GP_{33–632} Δ muc was transiently expressed in HEK293T cells with an N-terminal haemagglutinin tag for purification by affinity chromatography. GP was natively deglycosylated using PNGaseF and complexed with a SeMet-containing KZ52 antibody. The resulting trimeric GP-Fab complex was purified by size exclusion chromatography. Crystals were grown in 8.5% (w/v) PEG 10,000, 0.1 M Tris-HCl pH 8.5, 0.4 M sodium acetate and 10% (v/v) PEG 200 and diffracted to 3.4-Å resolution. Phases were obtained by combination of the Fab selenium anomalous signal and molecular replacement using an independently determined structure of the uncomplexed KZ52 Fab as a search model. Interpretable electron density maps with clear secondary structural elements and solvent boundaries were obtained after non-crystallographic symmetry (NCS)-averaged density modification. The structure was refined to final R_{work} and R_{free} values of 26.2% and 30.2%, respectively.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 February; accepted 14 May 2008.

- Sanchez, A. *et al.* In *Fields Virology* (eds Knipe, D. M. & Howley, P. M.) 1279–1304 (Lippincott, Williams & Wilkins, Philadelphia, 2001).
- Johnson, K. M., Lange, J. V., Webb, P. A. & Murphy, F. A. Isolation and characterization of a new virus causing acute hemorrhagic fever in Zaire. *Lancet* **1**, 569–571 (1977).
- Alsop, Z. Ebola outbreak in Uganda 'atypical', say experts. *Lancet* **370**, 2085 (2007).
- World Health Organization. Ebola hemorrhagic fever in Uganda. http://www.who.int/csr/don/2007_11_30a/en/index.html (2007).
- Peters, C. J., Sanchez, A., Rollin, P. E., Ksiazek, T. G. & Murphy, G. A. In *Fields Virology* Vol. 1, 1161–1176 (eds Fields, B. N., Knipe, D. M. & Howley, P. M.) (Lippincott-Raven Press, Philadelphia, 1996).
- Baize, S. *et al.* Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. *Nature Med.* **5**, 423–426 (1999).
- Ksiazek, T. G. *et al.* Clinical virology of Ebola hemorrhagic fever (EHF): virus, virus antigen, and IgG and IgM antibody findings among EHF patients in Kikwit, Democratic Republic of the Congo, 1995. *J. Infect. Dis.* **179** (suppl. 1), S177–S187 (1999).
- Leroy, E. M. *et al.* Human asymptomatic Ebola infection and strong inflammatory response. *Lancet* **355**, 2210–2215 (2000).
- Sanchez, A. *et al.* Analysis of human peripheral blood samples from fatal and nonfatal cases of Ebola (Sudan) hemorrhagic fever: cellular responses, virus load, and nitric oxide levels. *J. Virol.* **78**, 10370–10377 (2004).
- Hampton, T. Vaccines against Ebola and Marburg viruses show promise in primate studies. *J. Am. Med. Assoc.* **294**, 163–164 (2005).
- Jones, S. M. *et al.* Live attenuated recombinant vaccine protects nonhuman primates against Ebola and Marburg viruses. *Nature Med.* **11**, 786–790 (2005).
- Kobinger, G. P. *et al.* Chimpanzee adenovirus vaccine protects against Zaire Ebola virus. *Virology* **346**, 394–401 (2006).
- Sullivan, N. J. *et al.* Immune protection of nonhuman primates against Ebola virus with single low-dose adenovirus vectors encoding modified GPs. *PLoS Med.* **3**, e177 (2006).
- Sanchez, A., Trappier, S. G., Mahy, B. W., Peters, C. J. & Nichol, S. T. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc. Natl Acad. Sci. USA* **93**, 3602–3607 (1996).
- Sanchez, A. *et al.* Biochemical analysis of the secreted and virion glycoproteins of Ebola virus. *J. Virol.* **72**, 6442–6447 (1998).
- Volchkov, V. E., Feldmann, H., Volchkova, V. A. & Klenk, H. D. Processing of the Ebola virus glycoprotein by the proprotein convertase furin. *Proc. Natl Acad. Sci. USA* **95**, 5762–5767 (1998).
- Jeffers, S. A., Sanders, D. A. & Sanchez, A. Covalent modifications of the Ebola virus glycoprotein. *J. Virol.* **76**, 12463–12472 (2002).
- Feldmann, H., Klenk, H. D. & Sanchez, A. Molecular biology and evolution of filoviruses. *Arch. Virol.*, Suppl. 7, 81–100 (1993).
- Takada, A. *et al.* A system for functional analysis of Ebola virus glycoprotein. *Proc. Natl Acad. Sci. USA* **94**, 14764–14769 (1997).
- Wool-Lewis, R. J. & Bates, P. Characterization of Ebola virus entry by using pseudotyped viruses: identification of receptor-deficient cell lines. *J. Virol.* **72**, 3155–3160 (1998).
- Sanchez, A. Analysis of filovirus entry into Vero E6 cells, using inhibitors of endocytosis, endosomal acidification, structural integrity, and cathepsin (B and L) activity. *J. Infect. Dis.* **196** (suppl. 2), S251–S258 (2007).
- Chandran, K., Sullivan, N. J., Felbor, U., Whelan, S. P. & Cunningham, J. M. Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science* **308**, 1643–1645 (2005).
- Kaletsky, R. L., Simmons, G. & Bates, P. Proteolysis of the Ebola virus glycoproteins enhances virus binding and infectivity. *J. Virol.* **81**, 13378–13384 (2007).
- Schornberg, K. *et al.* Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *J. Virol.* **80**, 4174–4178 (2006).
- Alvarez, C. P. *et al.* C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. *J. Virol.* **76**, 6841–6844 (2002).
- Takada, A. *et al.* Human macrophage C-type lectin specific for galactose and N-acetylgalactosamine promotes filovirus entry. *J. Virol.* **78**, 2943–2947 (2004).
- Takada, A. *et al.* Downregulation of beta1 integrins by Ebola virus glycoprotein: implication for virus entry. *Virology* **278**, 20–26 (2000).
- Chan, S. Y. *et al.* Folate receptor-alpha is a cofactor for cellular entry by Marburg and Ebola viruses. *Cell* **106**, 117–126 (2001).
- Shimojima, M. *et al.* Tyro3 family-mediated cell entry of Ebola and Marburg viruses. *J. Virol.* **80**, 10109–10116 (2006).
- Malashkevich, V. N. *et al.* Core structure of the envelope glycoprotein GP2 from Ebola virus at 1.9-Å resolution. *Proc. Natl Acad. Sci. USA* **96**, 2662–2667 (1999).
- Weissenhorn, W., Carfi, A., Lee, K. H., Skehel, J. J. & Wiley, D. C. Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. *Mol. Cell* **2**, 605–616 (1998).
- Kuhn, J. H. *et al.* Conserved receptor-binding domains of Lake Victoria Marburgvirus and Zaire Ebolavirus bind a common receptor. *J. Biol. Chem.* **281**, 15951–15958 (2006).
- Yin, H. S., Wen, X., Paterson, R. G., Lamb, R. A. & Jardetzky, T. S. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. *Nature* **439**, 38–44 (2006).
- Delos, S. E. *et al.* Cysteines flanking the internal fusion peptide are required for the avian sarcoma/leukosis virus glycoprotein to mediate the lipid mixing stage of fusion with high efficiency. *J. Virol.* **82**, 3131–3134 (2008).
- Delos, S. E. & White, J. M. Critical role for the cysteines flanking the internal fusion peptide of avian sarcoma/leukosis virus envelope glycoprotein. *J. Virol.* **74**, 9738–9741 (2000).
- Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37–43 (1994).
- Wilson, I. A., Skehel, J. J. & Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366–373 (1981).
- Fass, D., Harrison, S. C. & Kim, P. S. Retrovirus envelope domain at 1.7 angstrom resolution. *Nature Struct. Biol.* **3**, 465–469 (1996).
- Rosenthal, P. B. *et al.* Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. *Nature* **396**, 92–96 (1998).
- Maruyama, T. *et al.* Ebola virus can be effectively neutralized by antibody produced in natural human infection. *J. Virol.* **73**, 6024–6030 (1999).
- Parren, P. W. H. I., Geisbert, T. W., Maruyama, T., Jahrling, P. B. & Burton, D. R. Pre- and post-exposure prophylaxis of Ebola virus infection in an animal model by passive transfer of a neutralizing human antibody. *J. Virol.* **76**, 6408–6412 (2002).
- Oswald, W. B. *et al.* Neutralizing antibody fails to impact the course of Ebola virus infection in monkeys. *PLoS Pathog.* **3**, 62–66 (2007).
- Maruyama, T. *et al.* Recombinant human monoclonal antibodies to Ebola virus. *J. Infect. Dis.* **179** (suppl. 1), S235–S239 (1999).
- Ritchie, G. E. The glycosylation of viral envelope glycoproteins and the effect of glycosidase inhibitors on virus replication and glycoprotein properties. DPhil thesis, Univ. Oxford (2005).
- Kwong, P. D. *et al.* Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648–659 (1998).
- Wyatt, R. *et al.* The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **393**, 705–711 (1998).
- Szakonyi, G. *et al.* Structure of the Epstein-Barr virus major envelope glycoprotein. *Nat. Struct. Mol. Biol.* **13**, 996–1001 (2006).
- Brindley, M. A. *et al.* Ebola virus glycoprotein 1: identification of residues important for binding and postbinding events. *J. Virol.* **81**, 7702–7709 (2007).
- Manicassamy, B., Wang, J., Jiang, H. & Rong, L. Comprehensive analysis of Ebola virus GP1 in viral entry. *J. Virol.* **79**, 4793–4805 (2005).
- Mpanju, O. M., Townner, J. S., Dover, J. E., Nichol, S. T. & Wilson, C. A. Identification of two amino acid residues on Ebola virus glycoprotein 1 critical for cell entry. *Virus Res.* **121**, 205–214 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Kimberlin (The Scripps Research Institute), the staff at the Advanced Light Source beamlines 8.3.1, 5.0.2, 4.2.2, 8.2.2 and 12.3.1 and the Stanford Synchrotron Radiation Laboratory beamlines 11-1 and 9-2 for data collection support, and D. Abelson and members of the Ollmann Saphire Laboratory for assistance, comments and suggestions. The Advanced Light Source and Stanford Synchrotron Radiation Laboratory are national user facilities operated on behalf of the US Department of Energy. We also thank A. Olson and J. Huntton for generation of tangible molecular models for analysis of receptor-binding surfaces. E.O.S. and D.R.B. are funded by the US National Institutes of Health, and E.O.S. and J.E.L. are supported by a Career Award from the Burroughs Wellcome Fund and a fellowship from the Canadian Institutes of Health Research, respectively. This is manuscript no. 19375 from The Scripps Research Institute.

Author Contributions and E.O.S. designed the initial GP constructs, and prepared GP and the initial GP–KZ52 crystallization screening experiments. J.E.L. designed the GP mutants, and purified, crystallized and determined the GP–KZ52 structure in the laboratory of E.O.S. A.J.H. expressed native and SeMet-incorporated IgG KZ52. W.B.O. performed infectivity and KZ52 neutralization studies, both in the laboratory of D.R.B. J.E.L. and E.O.S. wrote the manuscript.

Author Information Atomic coordinates and structure factors for the reported crystal structure have been deposited in the Protein Data Bank under accession number 3CSY. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to E.O.S. (erica@scripps.edu).

METHODS

Expression of EBOV GP and antibody. Zaire ebolavirus (Mayinga strain) glycoprotein DNA was codon-optimized, whole-gene synthesized (Blue Heron Biotech) and cloned into the pDISPLAY vector (Invitrogen) with a stop codon introduced before the internal transmembrane segment of the vector. GP is composed of more than 50% oligosaccharides by weight, primarily because of a heavily glycosylated mucin-like domain that is non-essential for cellular attachment or entry⁵¹. In an effort to increase sample homogeneity and solubility, deletion of the mucin-like and transmembrane domains (residues 312–462 and 633–676) and mutations of two N-linked glycosylation sites (T42V and T230V) were created by overlap PCR and the QuikChange II site-directed mutagenesis kit, respectively. Large-scale expression of T42V/T230V GP_{33–632}Δmuc was performed using HEK293T cells transfected by standard calcium phosphate precipitation in ten-layer CellStacks (Corning). The DNA–calcium phosphate mixture was added to 70% confluent cells grown in DMEM plus 1 × penicillin–streptomycin and 5% (v/v) fetal bovine serum. The supernatant was harvested four days post-transfection, concentrated using a Centrimate tangential flow system, purified by anti-HA immunoaffinity chromatography and natively deglycosylated with PNGaseF in PBS supplemented with 1.5 M urea at 37 °C.

Recombinant IgG1 KZ52 was expressed in Chinese hamster ovary (CHO-K1) cells in glutamine-free custom-formulated Glasgow minimum essential medium (GMEM) (MediaTech Cellgro). For large-scale tissue culture, media was supplemented with 3.5% (v/v) Ultra Low Bovine IgG Fetal Bovine Serum (Invitrogen) and grown in ten-layer CellStacks (Corning). The supernatants were centrifuged, sterile-filtered, purified over Protein A affinity matrix (GE Healthcare) and eluted in 0.1 M citric acid, pH 3.0. The pH of the eluted antibody was immediately brought to neutrality by the addition of 2 M Tris pH 9.0 and subsequently dialysed against phosphate-buffered saline (PBS, pH 7.4). Antibody concentration was determined by absorbance at 280 nm. Purity (>98%) and concentrations were confirmed by SDS–polyacrylamide gel electrophoresis. Selenomethionine-incorporated KZ52 was produced similarly, except cysteine- and methionine-free Roswell Park Memorial Institute (RPMI) (supplemented with 60 mg ml^{−1} L-selenomethionine and 60 mg l^{−1} L-cysteine) media were used for growth in T-225 cm² tissue culture flasks (Corning). Supernatant was harvested every 4 days over 2 weeks and concentrated using the Centrimate system before Protein A purification. IgG1 KZ52 was cleaved by standard papain digestion, the released Fc and uncleaved IgG were removed by Protein A chromatography and the resulting Fab fragments were further purified by MonoS ion exchange chromatography.

Preparation and crystallization of the GP–antibody complex. To promote the formation of crystal contacts, purified GP was mixed with excess Fab KZ52 and separated on a Superdex-200 GL 10/300 equilibrated with 10 mM Tris–HCl pH 7.5 and 150 mM NaCl. Fractions containing trimeric GP–Fab KZ52 were concentrated to about 8 mg ml^{−1} and crystallized by hanging-drop vapour diffusion at 22 °C in 8.5% (w/v) PEG 10,000, 0.1 M Tris–HCl pH 8.5, 0.4 M sodium acetate and 10% (v/v) PEG 200. These crystals were gently cross-linked with glutaraldehyde and cryoprotected with sequential soaks of 10%, 20%, 30% and 40% (v/v) glycerol plus mother liquor before flash cooling in liquid nitrogen. Although more than 800 crystals were screened at seven synchrotron beamlines, only three crystals diffracted to 4.0-Å resolution or better. Two complete data sets were collected at the Advanced Light Source beamline 5.0.2 from SeMet-derivatized Fab KZ52–GP crystals. All data were indexed, integrated and scaled using dT²REK⁵² to 3.4- and 4.0-Å resolution, respectively (Supplementary Table 1). The T42V/T230V GP_{33–632}Δmuc–KZ52 crystals belong to space group H32, with one and one-third trimers in the asymmetric unit.

Structure determination. Using an independently determined, uncomplexed Fab KZ52 structure as a search model, Phaser⁵³ identified one clear Fab solution residing on a threefold crystallographic axis (*z* score = 9.3). A trimeric arrangement of KZ52 Fab was generated by applying crystallographic symmetry, and the trimer of antibodies was subsequently used in a molecular replacement search while keeping the original Fab KZ52 solution fixed. Phaser identified a clear unique hit for the trimer (*z* score = 10.4). A difference Fourier anomalous electron-density map was calculated and selenium peaks for all 20 SeMet residues were picked manually using Xfit⁵⁴. Phase calculation combining Fab KZ52 model phases with Se-SAD anomalous data from two crystals was performed in SHARP⁵⁵. Initial phases from SHARP⁵⁵ were improved by solvent flipping, histogram matching and 1,500 cycles of NCS averaging using DM⁵⁶. Strong electron density was seen for the central regions of GP1 and GP2 in all four monomers, whereas weaker density was associated with the outer regions of GP1. Using the DM-improved phases, Resolve⁵⁷ was used to initially trace poly-alanine fragments in the map. These starting fragments were extended manually using Coot (v. 0.3.3)⁵⁸. All β-strands and α-helices were real-space refined with torsional secondary structural restraints. Calculation of B-value sharpened electron-density

maps improved the resolution of side-chain electron density and was used to help build the model. Torsion-angle simulated-annealing refinement, starting at 5,000 K⁵⁹ with a maximum-likelihood amplitude target, NCS restraints and no σ-cutoffs, was used in CNS-SOLVE (v. 1.2)⁶⁰. A final round of refinement using riding hydrogens, grouped atomic displacement and TLS parameters was implemented using Phenix.refine⁶¹. Identification of proper sequence registry was confirmed by secondary structural predictions and by anchoring N-linked glycosylation sequons to associated electron density. This strategy proved particularly useful as the N-linked glycan at N257 allowed initial assignment of the residues in the glycan cap (residues 214–267). Moreover, alternate sequence registers were considered in several portions of the GP1 and GP2 subunits, but these models could be eliminated based on inspection of σ_A-weighted 2F_o – F_c and F_o – F_c electron density maps. Procheck⁶², MolProbity⁶³ and composite omit electron-density maps (Supplementary Fig. 10) were used to validate the quality of the model. Although the protein used for crystallization was treated with protein N-glycosidase F (PNGaseF), two glycan chains at N257 on GP1 and N563 on GP2, resistant to deglycosylation, were retained. Electron density is visible for the chitobiose core and up to three mannose residues of these glycans (Supplementary Fig. 10). The pdb-care^{64,65} and carp⁶⁶ servers were used to detect discrepancies in glycan connectivity, nomenclature and glycosidic torsion angles. The final model contains EBOV GP residues 33–189, 214–278, 299–310 and 502–599, and Fab KZ52 residues 1–228 (heavy chain) and 1–211 (light chain). Electron density is missing for GP residues 190–213, 311–312, 464–501 and 600–632. Weak or discontinuous electron density is observed in the outer regions of the glycan cap (residues 268–278 and 299–310) and the loop containing the GP1–GP2 disulphide bridge (residues 49–56); these regions are tentatively assigned as poly-alanine fragments. Interestingly, a small chain-like piece of electron density is visible near N512 in GP2, but is not sufficiently resolved for identification. Final refinement statistics and analysis are presented in Supplementary Table 1. Molecular surface and ribbon diagrams were generated using MacPyMOL⁶⁷. For Fig. 4, N-linked complex-type glycans were modelled using the GlyProt server⁶⁸. Atomic coordinates and structure factors are deposited in the RCSB Protein Data Bank under accession number 3CSY.

Methods describing the assays for EBOV GP infectivity and KZ52 recognition at low pH are described in the Supplementary Methods.

51. Yang, Z. Y. *et al.* Identification of the Ebola virus glycoprotein as the main viral determinant of vascular cell cytotoxicity and injury. *Nature Med.* **6**, 886–889 (2000).
52. Pflugrath, J. W. The finer things in X-ray diffraction data collection. *Acta Crystallogr. D* **55**, 1718–1725 (1999).
53. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
54. McRee, D. M. *Practical Protein Crystallography* (Academic Press, San Diego, 1993).
55. de La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494 (1997).
56. Cowtan, K. Dm: an automated procedure for phase improvement by density modification. *Joint CCP4 ESF-EACBM News. Protein Crystallogr.* **31**, 34–38 (1994).
57. Terwilliger, T. C. Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D* **59**, 38–44 (2003).
58. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
59. Brünger, A. T., Krukowski, A. & Erickson, J. W. Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr. A* **47**, 195–204 (1990).
60. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
61. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
62. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
63. Davis, I. W., Murray, L. W., Richardson, J. S. & Richardson, D. C. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32**, W615–W619 (2004).
64. Lutteke, T., Frank, M. & von der Lieth, C. W. Data mining the protein data bank: automatic detection and assignment of carbohydrate structures. *Carbohydr. Res.* **339**, 1015–1020 (2004).
65. Lutteke, T. & von der Lieth, C. W. pdb-care (PDB Carbohydrate REsidue check): a program to support annotation of complex carbohydrate structures in PDB files. *BMC Bioinformatics* **5**, 69 (2004).
66. Lutteke, T., Frank, M. & von der Lieth, C. W. Carbohydrate Structure Suite (CSS): analysis of carbohydrate 3D structures derived from the PDB. *Nucleic Acids Res.* **33**, D242–D246 (2005).
67. DeLano, W. L. The PyMol Molecular Graphics System (DeLano Scientific, Palo Alto, California, 2002).
68. Bohne-Lang, A. & von der Lieth, C. W. GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res.* **33**, W214–W219 (2005).

Crystal structure of the ligand-free G-protein-coupled receptor opsin

Jung Hee Park^{1*}, Patrick Scheerer^{1*}, Klaus Peter Hofmann^{1,2}, Hui-Woog Choe^{1,3} & Oliver Peter Ernst¹

In the G-protein-coupled receptor (GPCR) rhodopsin, the inactivating ligand 11-*cis*-retinal is bound in the seven-transmembrane helix (TM) bundle and is *cis/trans* isomerized by light to form active metarhodopsin II. With metarhodopsin II decay, all-*trans*-retinal is released, and opsin is reloaded with new 11-*cis*-retinal. Here we present the crystal structure of ligand-free native opsin from bovine retinal rod cells at 2.9 ångström (Å) resolution. Compared to rhodopsin, opsin shows prominent structural changes in the conserved E(D)RY and NPxxY(x)_{5,6}F regions and in TM5–TM7. At the cytoplasmic side, TM6 is tilted outwards by 6–7 Å, whereas the helix structure of TM5 is more elongated and close to TM6. These structural changes, some of which were attributed to an active GPCR state, reorganize the empty retinal-binding pocket to disclose two openings that may serve the entry and exit of retinal. The opsin structure sheds new light on ligand binding to GPCRs and on GPCR activation.

GPCRs, or seven-transmembrane receptors, constitute one of the largest superfamilies in the human genome¹. As cell-surface receptors, GPCRs mediate responses to visual, olfactory, hormonal, neurotransmitter and other signals. Binding of small extracellular ligands modulates the capacity of GPCRs to catalyse GDP/GTP exchange in heterotrimeric guanine-nucleotide-binding proteins (G proteins) inside the cell. Rhodopsin is the eponym of the largest GPCR family¹; it is used in scotopic vision and is found in high concentrations in rod photoreceptor cells². Light-sensitive rhodopsin consists of the apoprotein opsin and the chromophore 11-*cis*-retinal which is covalently bound to Lys 296 in TM7 by a protonated Schiff base. As a strong inverse agonist, 11-*cis*-retinal within the ligand-binding pocket stabilizes the inactive receptor state until photon absorption causes retinal *cis/trans* isomerization. Thereby the agonist all-*trans*-retinal is formed *in situ*, which drives the receptor to the active state. Light-induced conformational changes are reflected in distinct photointermediates and culminate in active metarhodopsin II containing a deprotonated retinylidene Schiff base³. Known structural features that change with metarhodopsin II formation include breakage of constraints within the conserved E(D)RY and NPxxY(x)_{5,6}F motifs and an outward tilt of the cytoplasmic part of TM6 (refs 2–8).

The G-protein activation phase is terminated by the interaction of metarhodopsin II with rhodopsin kinase and arrestin². Subsequently, the retinal Schiff base is hydrolysed and the photolysed all-*trans*-retinal is released from its binding site. Regeneration of the light-sensitive rhodopsin ground state requires the supply of new 11-*cis*-retinal through the so-called retinoid cycle^{9,10}. As long as the retinal-binding site is empty and the opsin apoprotein waits for reconstitution with 11-*cis*-retinal, one experiences a reduction of visual sensitivity⁹. The reduced sensitivity reflects a background level of active G protein, which is thought to arise from a catalytically active form of opsin⁹. An active conformation of opsin, Ops*, was also found *in vitro* by infrared spectroscopy¹¹.

Several crystal structures of the 11-*cis*-retinal-bound rhodopsin ground state^{12–16}, of early photointermediates¹⁷ and of a retinal Schiff base deprotonated state¹⁵ have been solved. A close similarity

is found between all these structures, which also applies to the recently solved structures of a second GPCR, the β_2 -adrenergic receptor^{18–20}. This GPCR was stabilized in the inactive state by binding of a partial inverse agonist. Protein engineering^{18,19} or binding of an antibody fragment (Fab)²⁰ facilitated crystallization of the β_2 -adrenergic receptor.

Ligand-free GPCRs are difficult to purify and have so far resisted attempts at their crystallization. In the case of opsin, it is known that purified opsin is unstable and tends to aggregate. An opsin mutant with increased thermal stability could be purified and formed rhodopsin on reconstitution with 11-*cis*-retinal²¹. This mutant rhodopsin contained a disulphide bond tethering the amino-terminal extracellular domain to the extracellular loop connecting TM6 and TM7, and allowed the first crystallization and structure determination of a recombinant GPCR¹⁶. Here we describe the X-ray structure of native opsin, a GPCR in its empty state, to gain insight into the uptake and release of retinal and ligand binding to GPCRs in general.

Structure of opsin

To obtain crystals of native opsin, we took a different route and optimized the extraction of bovine opsin from rod cell disc membranes, inspired by a method developed earlier for selective extraction of rhodopsin²². Avoiding further purification steps enabled us to obtain crystals of opsin in its native state without any modification of the protein that may cause structural distortions. For data collection, structure determination and refinement statistics, see Methods and Supplementary Table 1.

The rhombohedral crystals contain two opsin molecules (A and B) and six β -D-octylglucoside molecules in the asymmetric unit. The final opsin model includes residues 1–326, two oligosaccharides attached to Asn 2 and Asn 15, and a palmitoylate chain attached to Cys 322 (Fig. 1). The carboxy-terminal cytoplasmic domain of opsin (residues 327–348) is partially disordered and is not determined owing to weak electron density, whereas the transmembrane region is clearly resolved. The overall topology is similar to that of rhodopsin, with seven transmembrane-spanning helical segments

¹Institut für Medizinische Physik und Biophysik (CC2), Charité-Universitätsmedizin Berlin, Charitéplatz 1, D-10117 Berlin, Germany. ²Zentrum für Biophysik und Bioinformatik, Humboldt-Universität zu Berlin, Invalidenstrasse 42, D-10115 Berlin, Germany. ³Department of Chemistry, College of Natural Science, Chonbuk National University, 561-756 Chonju, South Korea.

*These authors contributed equally to this work.

connected by three cytoplasmic loops (C1 to C3) and three extracellular loops (E1 to E3), and a cytoplasmic helix (H8) adjacent to TM7 running along the membrane surface. The two monomers in the asymmetric unit are arranged as a dimer with TM1 and H8 as an interface (Fig. 1). Similar arrangements were found in two-dimensional and three-dimensional crystals of rhodopsin^{15,23,24}. In disc membranes, rhodopsin dimers can form as observed by atomic force microscopy, and TM4–TM5 was suggested as a dimer interface². In the structures of rhodopsin and opsin, a disulphide bond between Cys 110 in TM3 and Cys 187 in loop E2 is present. Palmitoylation of opsin was clearly resolved for Cys 322 but not for Cys 323, presumably owing to the high flexibility of the fatty acid chain at Cys 323. Only ten water molecules were modelled in the refinement due to weak electron density.

Within the helix bundle, opsin shows only small changes relative to rhodopsin for TM1–TM4 (Fig. 2), reminiscent of the proposed stable TM1–TM4 core structure that does not change much on receptor activation²⁵. The most prominent change in this core is the formation of a short helical turn in loop C1 of opsin. Larger changes are observed for TM5–TM7, especially dominant at the cytoplasmic ends of these helices, causing rearrangement of loops C2 and C3 and the kink region between TM7 and H8. The N-terminal part of loop C3 adopts an α -helical structure. Therefore, TM5 is 1.5–2.5 helical turns longer in opsin than in rhodopsin, depending on the reference model. TM5 also slightly tilts in opsin, hence the cytoplasmic TM5 end is shifted by 2–3 Å towards TM6. In both opsin and rhodopsin, a kink induced by Pro 267 divides TM6 into two segments. The whole TM6 in opsin is slightly displaced away from the centre of the helix bundle. At the cytoplasmic end, TM6 is tilted outward from the helix bundle by 6–7 Å with Trp 265 as a pivot point. The cytoplasmic segment of TM6 therefore comes closer to TM5 so that it runs almost parallel to TM5. A rigid body movement of the cytoplasmic segment of TM6 was measured by electron paramagnetic resonance spectroscopy as a concomitant event during the light-induced activation of rhodopsin^{4,7}. The large outward tilt in the opsin structure of TM6 would thus indicate an active or partially active opsin conformation, in agreement with the evidence for Ops* from electrophysiological, spectroscopic and mutational analysis^{9,11,26}.

In rhodopsin, a salt bridge occurs between Arg 135 (Arg^{3,50}; superscript refers to the generalized Ballesteros–Weinstein numbering

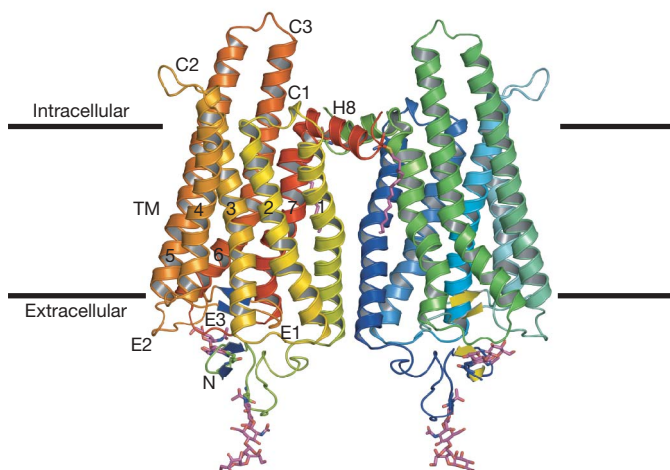


Figure 1 | Overall fold of opsin. The two monomers in the asymmetric unit are viewed from within the membrane. Seven transmembrane helices (TM1–TM7), which are connected by extracellular (E) and cytoplasmic (C) loops, are followed by a cytoplasmic helix (H8). The amino terminus is labelled 'N'. Also shown are N-linked oligosaccharides attached to Asn 2 and Asn 15 and a palmitoylate chain at Cys 322 (at the end of H8). Twenty-two C-terminal amino acids are not included in the model.

system²⁷) in TM3—which is part of the conserved E(D)RY motif in GPCRs¹²—and Glu 247 (Glu^{6,30}) in TM6. Together with Glu 134 and Thr 251, they form a more extended hydrogen-bonded network, the so-called 'ionic lock'²⁸, which stabilizes TM6 within the helical bundle^{12–14} (Fig. 3a). In opsin, the ionic lock is broken and the rearranged cytoplasmic ends of TM5 and TM6 are locked by two new

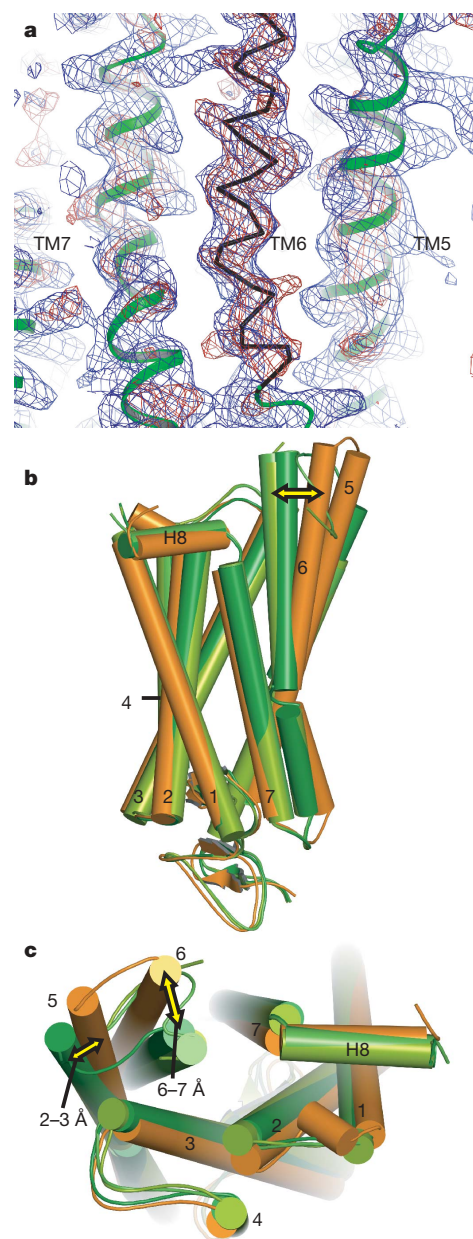


Figure 2 | Structure of opsin and comparison with rhodopsin. **a**, $2F_o - F_c$ and $F_o - F_c$ electron density maps contoured at 1.0σ (blue mesh) and 2.0σ (red mesh), respectively. Electron density maps were calculated using data to 4.0 Å resolution and initial molecular replacement phases, which were obtained using a truncated rhodopsin model (PDB accession 1U19) lacking both extracellular and cytoplasmic regions and TM6. The truncated initial model is shown as a green cartoon and the first C α trace of TM6 is shown as a black ribbon. **b**, **c**, Comparison of helix orientations: superposition of C α traces of opsin (orange) and different rhodopsin structures (PDB accessions 1F88 (lime-green); 1U19 (green); and 1GZM (forest-green)). Helices are shown as rods; loop regions were smoothed for clarity. The side view (**b**) and the view from the cytoplasmic side (**c**) are shown. The yellow double arrows indicate different positions of cytoplasmic segment of TM6, with Trp 265 as the pivot point of movement.

interactions (Fig. 3b). Arg 135, which is released from Glu 134 and Glu 247, engages with Tyr 223 to tether TM5. Glu 247 flips over to interact in an extended hydrogen-bonded network with Lys 231 in TM5 and Thr 251 in TM6 to stabilize the TM5 and TM6 pair at the cytoplasmic side. Glu 134 is now positioned towards TM2 and TM4 and does not seem to undergo any specific interaction (not shown). In contrast, crystal structures of early photoproducts of rhodopsin¹⁷, a late photoproduct with deprotonated retinal Schiff base¹⁵ and the β_2 -adrenergic receptor bound to a partial inverse agonist^{18–20} show an Arg^{3,50}–Glu^{6,30} distance that is only slightly larger than in rhodopsin.

Compared to rhodopsin, the length and orientation of H8 remain unchanged in opsin. In opsin, TM7 deviates most from a regular α -helical structure (as in rhodopsin) and shows an even more distinct S-shape. TM7 contains the conserved NPxxY(x)_{5,6}F motif^{6,12}, which links TM7 and H8 by an electrostatic interaction between the aromatic side chains of Tyr 306 (Tyr^{7,53}) and Phe 313 (Phe^{7,60}) (Fig. 4a). The cytoplasmic end of TM7 reorganizes in opsin such that Tyr 306 rotates to face into the helix bundle (Fig. 4b). It thereby blocks TM6 from moving back towards TM3 to adopt the conformation of rhodopsin.

Rhodopsin contains a compact extracellular domain, the so-called 'retinal plug'²⁹, including two antiparallel β -sheets in the N terminus (strands β 1 and β 2) and loop E2 (strands β 3 and β 4). Loop E2 folds back into the protein to contribute residues Ser 186 to Ile 189 of strand β 4 to the retinal-binding site. This whole retinal plug structure with the antiparallel β -sheets is maintained in opsin (Figs 1 and 2).

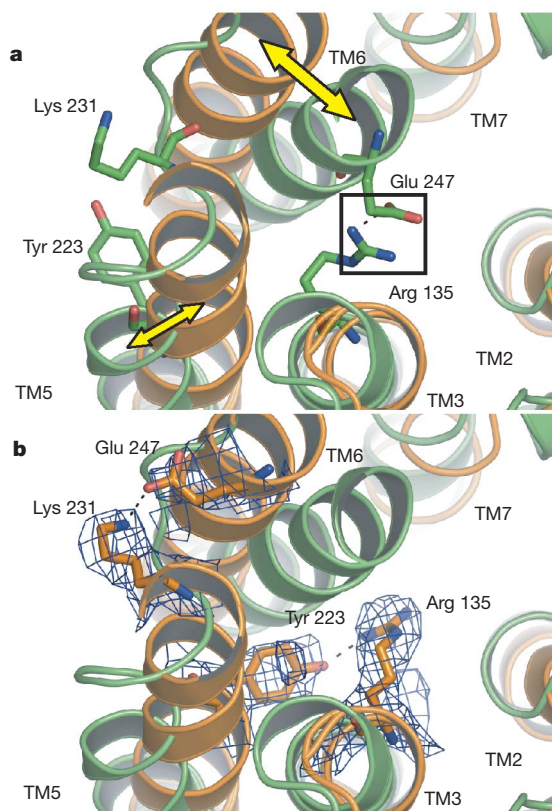


Figure 3 | Superposition of conserved E(D)RY regions of rhodopsin and opsin. Rhodopsin (green; PDB accession 1U19) and opsin (orange) viewed from the cytoplasmic side. The yellow double arrows indicate different positions of TM5 and TM6. **a**, An extended hydrogen-bonded network termed ionic lock between TM3 and TM6 (Arg 135–Glu 247; also including Glu 134 in TM3 and Thr 251 in TM6, which are not shown) is present in rhodopsin, which constrains TM6 inside the helix bundle. **b**, In opsin, a different set of interactions (Lys 231–Glu 247 and Tyr 223–Arg 135) is formed to stabilize the cytoplasmic TM5 and TM6 pair. The $2F_o - F_c$ electron density map is contoured at 1.5σ (blue mesh).

Within the retinal-binding pocket, the lack of the interacting prosthetic group 11-*cis*-retinal causes a few distinct structural alterations in the retinal-binding pocket (Fig. 5). In rhodopsin, the β -ionone ring of retinal is kept in place by Gly 121 and Glu 122 in TM3, Met 207, His 211 and Phe 212 in TM5, and Phe 261, Trp 265, Tyr 268 and Ala 269 in TM6 (ref. 12). The indole side chain of the highly conserved Trp 265 (Trp^{6,48}) is in van der Waals contact with the β -ionone ring of retinal. Part of the space occupied by the β -ionone ring is filled in opsin with the side chain of Trp 265, concomitant with a movement of Phe 261 towards TM5. Trp 265 is therefore in apposition to Glu 122 (not shown in Fig. 5), with probably enough space between these residues to accommodate a water molecule. Coupled changes in the rotamer states of Trp 265 and the β -ionone ring, or of Trp^{6,48} and Phe^{6,52} for GPCRs in general, were predicted to enable activating outward movement of the cytoplasmic end of TM6 (refs 30 and 31). The opsin and rhodopsin structures do not differ by such a rotamer change of Trp 265 but show different locations for the cytoplasmic end of TM6. However, lack of a Trp rotamer change may be due to the empty retinal-binding pocket.

In rhodopsin, retinal is held along the polyene chain by residues Ala 117, Thr 118, Gly 120 and Gly 121 in TM3, by Tyr 268 in TM6 and by Ser 186–Ile 189 and Tyr 191 in loop E2. In opsin, the extracellular part of TM3 and loop E2 are slightly moved away from helices TM5–TM7. Thereby, the retinal-binding pocket becomes wider towards the retinal attachment site Lys 296 and probably allows harbouring of extra water molecules. The fact that very low electron density is observed for the side chain of Lys 296 suggests that it is very flexible. Thus, the ϵ -amino group of Lys 296 does not seem to be involved in a salt bridge with Glu 113 in TM3 or with Glu 181 in loop E2. Glu 113 is the retinal Schiff base counterion in the rhodopsin dark state; Glu 181 was proposed to be part of a complex counterion which forms in metarhodopsin I, the inactive precursor of metarhodopsin II (ref. 32). From studies on opsin mutants it was concluded that breakage of the salt bridge between Lys 296 and Glu 113, together with sufficiently low pH, is mandatory for the activation of opsin²⁶. Residues Tyr 43, Met 44 and Leu 47 in TM1, Phe 91 and Thr 94 in TM2, and Phe 293 in TM7 confine the retinal pocket around Lys 296 in rhodopsin. These residues remain unaltered in opsin except for Phe 293, in which the side chain is rotated upward by $\sim 120^\circ$ and the residue is somewhat moved away from TM1 owing to a movement of the backbone. Also, the extracellular end of TM5 and adjacent loop E2 are shifted slightly towards TM6. As mentioned previously, the retinal-binding pocket is tightly sealed by the retinal plug. Consistent with biochemical studies²¹, this argues against the idea that the plug opens in opsin to allow uptake or release of retinal from the extracellular side.

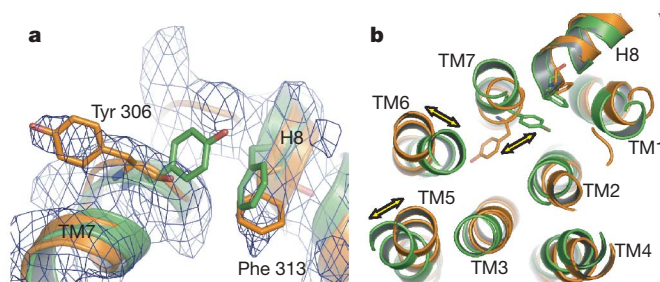


Figure 4 | Superposition of conserved NPxxY(x)_{5,6}F regions of rhodopsin and opsin. Rhodopsin (green; PDB accession 1U19; **a**) and opsin (orange; **b**) viewed from the cytoplasmic side. The yellow double arrows indicate differences in positions of TM5, TM6 and Tyr 306. **a**, The interaction between Tyr 306 and Phe 313 is only observed in rhodopsin. The $2F_o - F_c$ electron density map for opsin is contoured at 1.5σ (blue mesh). **b**, In opsin, Tyr 306 and Phe 313 do not interact because Tyr 306 is oriented into the helix bundle.

Retinal exchange

In contrast to rhodopsin, the opsin structure shows two openings of the retinal-binding pocket: one between the extracellular ends of TM5 and TM6, and another between TM1 and TM7 (Fig. 6a). Owing to movement of TM6, an opening is formed by the membrane-facing residues Ile 205 and Phe 208 in TM5, and residues Phe 273 and Phe 276 in TM6 (Fig. 6a and inset 1). This hydrophobic hole, which is not observed in rhodopsin, is caused by small backbone movements and different side-chain orientations (Fig. 6b). Below this opening, the extracellular ends of TM5 and TM6 are held together by a weak hydrogen bond between Glu 201 in TM5 and Gln 279 in loop E3, next to TM6. Loop E3 shows a slightly different conformation in opsin compared to rhodopsin. The second, smaller opening arises from backbone alteration in TM7 in the location of the retinal-binding pocket and the corresponding clockwise $\sim 120^\circ$ rotation of Phe 293 relative to rhodopsin (Fig. 6a and inset 2). Thereby, a small hole between TM1 and TM7 is generated, which at the inner face of the helices can be sealed by rotamers of the Lys 296 side chain.

These two openings of the retinal-binding site hint at different retinal entrance and exit routes that were proposed for retinal channelling in opsin³³. The opening between TM5 and TM6 could be selective for uptake of 11-*cis*-retinal, as outlined in the Supplementary Discussion. Regeneration of rhodopsin from the opsin apoprotein can only occur when the retinal-binding site is freed from the photolysed all-*trans*-retinal. When neglecting a complex bypass—that is, partial storage of retinal in metarhodopsin III (refs 34 and 35)—the first step is the hydrolysis of the retinal Schiff base. A strong increase in intrinsic fluorescence^{33,36} indicates that the interaction between the retinal β -ionone ring and Trp 265 is consequently broken. It was postulated that all-*trans*-retinal exits via a route that is different from the entrance route, and remains bound to an exit site³³. The opening between TM1 and TM7 (Fig. 6a) provides a candidate for a release path for all-*trans*-retinal. A putative external binding site for retinal, which may also serve as an exit site, could be provided by the TM7 and H8 kink region. The retinal β -ionone ring in opsin could replace Tyr 306 (which interacts with Phe 313 in rhodopsin; Fig. 4a), and fix Tyr 306 inside the helical bundle. Because this would stabilize TM6 in the outward position, it would also explain the partial activity when all-*trans*-retinal is added and forms an adduct with opsin^{37–39}. Although the mechanistic aspects of these forms of opsin remain to be elucidated, their potential significance for vision makes them rewarding targets for future research. They may also

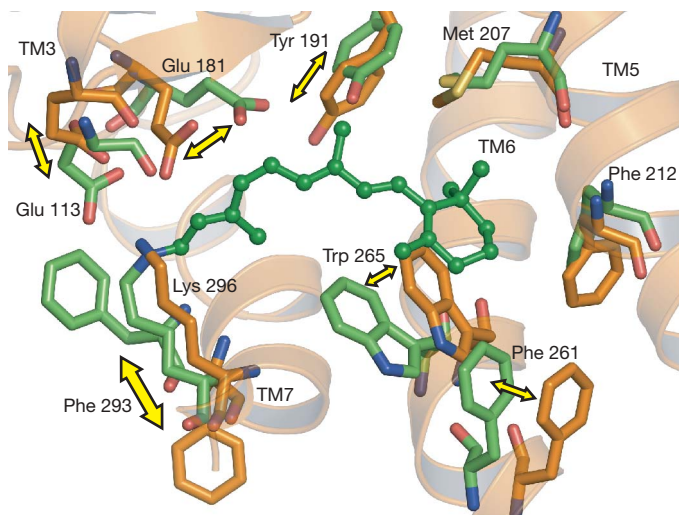


Figure 5 | Retinal-binding pocket. Residues of the retinal-binding pocket in rhodopsin (green; PDB accession 1U19) and opsin (orange) are rendered as sticks. The yellow double arrows indicate differences in side-chain positions. Retinal in rhodopsin is shown as a ball-and-stick model for reference. In opsin, Lys 296 is flexible and a rotamer was modelled for reference.

have a role in disorders in the regeneration pathway, which have been associated with different forms of blindness such as Oguchi's and Stargardt's disease⁹.

Comparison with other GPCRs

Rhodopsin differs from other GPCRs by its light-induced fast pathway from *cis/trans* retinal photoisomerization to active metarhodopsin II, and the associated formation and hydrolysis reactions of the retinal Schiff base. However, apart from this photochemical functional core, rhodopsin shows properties that also apply for GPCRs activated by diffusible ligands. GPCRs exist in multiple inactive and active conformations that are stabilized by ligands and/or G-protein⁴⁰ or pH^{26,41}. Different lines of evidence have shown that an active opsin conformation is present in the rod cell until fresh 11-*cis*-retinal is taken up to regenerate rhodopsin^{9,10,42}. The amount of active Ops* is low under cellular conditions, but can be much enhanced at low pH according to a pH-dependent Ops \rightleftharpoons Ops* equilibrium^{11,26}. Opsin, devoid of the strongly inactivating chromophore 11-*cis*-retinal, thus has properties of a classical GPCR.

The proposed uptake route of retinal between the extracellular ends of TM5 and TM6 (see Supplementary Information) may also be more general for GPCRs. For example, the amine ligands of the β_2 -adrenergic receptor contain similar structures to retinal, which include a hydrophobic moiety and a terminal amino group. In the structure of the carazolol-bound β_2 -adrenergic receptor there is a short helical segment in the middle of loop E2 that is constrained by an intraloop disulphide bridge¹⁸. Loop E2 is in addition tethered to TM3, making it probable that the aromatic carbazole ring of carazolol enters the ligand-binding pocket at the opposite side, between TM5 and TM6. Because many ligands of GPCRs are more hydrophilic than retinal, the observed α -helix in loop E2 and the disordered N terminus in the β_2 -adrenergic receptor could have a specific role, namely of capturing diffusible ligands and providing a wider opening to allow the more hydrophilic end of ligands to enter the binding site deeper in the receptor.

Relative to inactive 11-*cis*-retinal-bound rhodopsin, ligand-free opsin shows large changes at the G-protein coupling surface. These changes include the outwards tilt of TM6, the elongation and side-wards shift of TM5, breakage of the ionic lock between TM3 and TM6, a broken aromatic interaction between TM7 and H8, and new interactions stabilizing TM3 and TM5, and TM5 and TM6, respectively. Some of these features were predicted to be properties of the active, G-protein-interacting metarhodopsin II state^{2–8}.

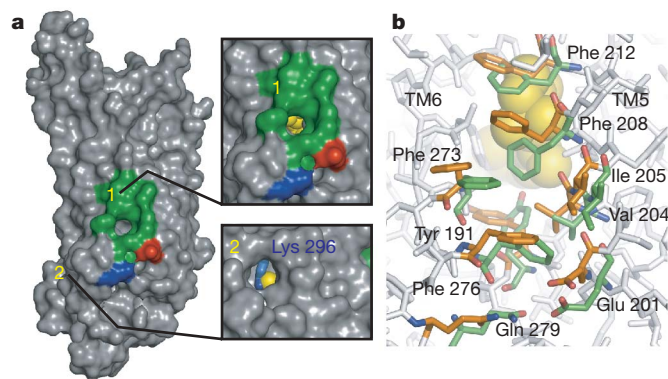


Figure 6 | Openings of the retinal-binding pocket in opsin. **a**, The side views of the opening between TM5 and TM6 (inset 1) and between TM1 and TM7 (inset 2) are shown. In the inset views, superimposed retinal from rhodopsin (PDB accession 1U19) is shown as yellow spheres for reference. No defined electron density was found for Lys 296. Lys 296 from superimposed rhodopsin (PDB accession 1U19) is shown in blue. **b**, Stick representations of the opening between TM5 and TM6 of opsin (orange) and rhodopsin (green; PDB accession 1U19) are shown. 11-*cis*-retinal is shown as yellow spheres.

Metarhodopsin II comprises different substates with different types and degrees of activity⁷. This could explain why only small structural changes were observed in a photoactivated intermediate of rhodopsin with deprotonated retinal Schiff base¹⁵. Although it is conceivable that we have crystallized the active form of opsin, Ops*, we do not know whether the features we have found comprise the whole structural arsenal that is required to activate the G protein. Only the structure of the complex between light-activated rhodopsin and the G protein or peptides derived from its receptor interacting domains will provide an answer to this question.

METHODS SUMMARY

Native opsin was prepared from bovine eyes and extracted from rod outer-segment disc membranes using β -D-octylglucopyranoside. Without further purification, solubilized opsin was crystallized by hanging-drop vapour diffusion in a mixture of ammonium sulphate and MES or sodium acetate buffer at pH 5.6. Crystals grown within 7 days were cryoprotected in 10% trehalose and frozen in liquid nitrogen for X-ray analysis at the synchrotron BESSY (Berlin, Germany). The opsin structure was solved by molecular replacement using several models of rhodopsin monomers (Protein Data Bank (PDB) accessions 1U19, 1GZM and 1HZX) without cytoplasmic and extracellular regions as initial search trials.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 March; accepted 9 May 2008.

Published online 18 June 2008.

- Lagerstrom, M. C. & Schioth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nature Rev. Drug Discov.* **7**, 339–357 (2008).
- Palczewski, K. G protein-coupled receptor rhodopsin. *Annu. Rev. Biochem.* **75**, 743–767 (2006).
- Okada, T., Ernst, O. P., Palczewski, K. & Hofmann, K. P. Activation of rhodopsin: new insights from structural and biochemical studies. *Trends Biochem. Sci.* **26**, 318–324 (2001).
- Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L. & Khorana, H. G. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* **274**, 768–770 (1996).
- Sheikh, S. P., Zvyaga, T. A., Lichtarge, O., Sakmar, T. P. & Bourne, H. R. Rhodopsin activation blocked by metal-ion-binding sites linking transmembrane helices C and F. *Nature* **383**, 347–350 (1996).
- Fritze, O. *et al.* Role of the conserved NPxxY(x)₅6F motif in the rhodopsin ground state and during activation. *Proc. Natl Acad. Sci. USA* **100**, 2290–2295 (2003).
- Knierim, B., Hofmann, K. P., Ernst, O. P. & Hubbell, W. L. Sequence of late molecular events in the activation of rhodopsin. *Proc. Natl Acad. Sci. USA* **104**, 20290–20295 (2007).
- Menon, S. T., Han, M. & Sakmar, T. P. Rhodopsin: structural basis of molecular physiology. *Physiol. Rev.* **81**, 1659–1688 (2001).
- Lamb, T. D. & Pugh, E. N. Jr. Dark adaptation and the retinoid cycle of vision. *Prog. Retin. Eye Res.* **23**, 307–380 (2004).
- McBee, J. K., Palczewski, K., Baehr, W. & Pepperberg, D. R. Confronting complexity: the interlink of phototransduction and retinoid metabolism in the vertebrate retina. *Prog. Retin. Eye Res.* **20**, 469–529 (2001).
- Vogel, R. & Siebert, F. Conformations of the active and inactive states of opsin. *J. Biol. Chem.* **276**, 38487–38493 (2001).
- Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739–745 (2000).
- Okada, T. *et al.* The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **342**, 571–583 (2004).
- Li, J., Edwards, P. C., Burghammer, M., Villa, C. & Schertler, G. F. Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **343**, 1409–1438 (2004).
- Salom, D. *et al.* Crystal structure of a photoactivated deprotonated intermediate of rhodopsin. *Proc. Natl Acad. Sci. USA* **103**, 16123–16128 (2006).
- Standfuss, J. *et al.* Crystal structure of a thermally stable rhodopsin mutant. *J. Mol. Biol.* **372**, 1179–1188 (2007).
- Nakamichi, H. & Okada, T. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc. Natl Acad. Sci. USA* **103**, 12729–12734 (2006).
- Cherezov, V. *et al.* High-resolution crystal structure of an engineered human β_2 -adrenergic G protein-coupled receptor. *Science* **318**, 1258–1265 (2007).
- Rosenbaum, D. M. *et al.* GPCR engineering yields high-resolution structural insights into β_2 -adrenergic receptor function. *Science* **318**, 1266–1273 (2007).
- Rasmussen, S. G. *et al.* Crystal structure of the human β_2 adrenergic G-protein-coupled receptor. *Nature* **450**, 383–387 (2007).
- Xie, G., Gross, A. K. & Oprian, D. D. An opsin mutant with increased thermal stability. *Biochemistry* **42**, 1995–2001 (2003).
- Okada, T., Takeda, K. & Kouyama, T. Highly selective separation of rhodopsin from bovine rod outer segment membranes using combination of divalent cation and alkyl(thio)glucoside. *Photochem. Photobiol.* **67**, 495–499 (1998).
- Krebs, A., Edwards, P. C., Villa, C., Li, J. & Schertler, G. F. The three-dimensional structure of bovine rhodopsin determined by electron cryomicroscopy. *J. Biol. Chem.* **278**, 50217–50225 (2003).
- Rupprecht, J. J., Mielke, T., Vogel, R., Villa, C. & Schertler, G. F. Electron crystallography reveals the structure of metarhodopsin I. *EMBO J.* **23**, 3609–3620 (2004).
- Liu, W., Eilers, M., Patel, A. B. & Smith, S. O. Helix packing moments reveal diversity and conservation in membrane protein structure. *J. Mol. Biol.* **337**, 713–729 (2004).
- Cohen, G. B., Oprian, D. D. & Robinson, P. R. Mechanism of activation and inactivation of opsin: role of Glu113 and Lys296. *Biochemistry* **31**, 12592–12601 (1992).
- Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G-protein coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).
- Ballesteros, J. A. *et al.* Activation of the β_2 -adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. *J. Biol. Chem.* **276**, 29171–29177 (2001).
- Filipek, S., Stenkamp, R. E., Teller, D. C. & Palczewski, K. G protein-coupled receptor rhodopsin: a prospectus. *Annu. Rev. Physiol.* **65**, 851–879 (2003).
- Crocker, E. *et al.* Location of Trp265 in metarhodopsin II: implications for the activation mechanism of the visual receptor rhodopsin. *J. Mol. Biol.* **357**, 163–172 (2006).
- Shi, L. *et al.* β_2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *J. Biol. Chem.* **277**, 40989–40996 (2002).
- Lüdeke, S. *et al.* The role of Glu181 in the photoactivation of rhodopsin. *J. Mol. Biol.* **353**, 345–356 (2005).
- Schädel, S. A. *et al.* Ligand channeling within a G-protein-coupled receptor. The entry and exit of retinals in native opsin. *J. Biol. Chem.* **278**, 24896–24903 (2003).
- Heck, M. *et al.* Signaling states of rhodopsin. Formation of the storage form, metarhodopsin III, from active metarhodopsin II. *J. Biol. Chem.* **278**, 3162–3169 (2003).
- Bartl, F. J. & Vogel, R. Structural and functional properties of metarhodopsin III: Recent spectroscopic studies on deactivation pathways of rhodopsin. *Phys. Chem. Chem. Phys.* **9**, 1648–1658 (2007).
- Farrens, D. L. & Khorana, H. G. Structure and function in rhodopsin. Measurement of the rate of metarhodopsin II decay by fluorescence spectroscopy. *J. Biol. Chem.* **270**, 5073–5076 (1995).
- Hofmann, K. P., Pulvermüller, A., Buczylo, J., Van Hooser, P. & Palczewski, K. The role of arrestin and retinoids in the regeneration pathway of rhodopsin. *J. Biol. Chem.* **267**, 15701–15706 (1992).
- Jäger, S., Palczewski, K. & Hofmann, K. P. Opsin/all-trans-retinal complex activates transducin by different mechanisms than photolyzed rhodopsin. *Biochemistry* **35**, 2901–2908 (1996).
- Sachs, K., Maretzki, D., Meyer, C. K. & Hofmann, K. P. Diffusible ligand all-trans-retinal activates opsin via a palmitoylation-dependent mechanism. *J. Biol. Chem.* **275**, 6189–6194 (2000).
- De Lean, A., Stadel, J. M. & Lefkowitz, R. J. A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled beta-adrenergic receptor. *J. Biol. Chem.* **255**, 7108–7117 (1980).
- Scheer, A., Fanelli, F., Costa, T., De Benedetti, P. G. & Cotecchia, S. The activation process of the α_{1B} -adrenergic receptor: potential role of protonation and hydrophobicity of a highly conserved aspartate. *Proc. Natl Acad. Sci. USA* **94**, 808–813 (1997).
- Kefalov, V. J., Crouch, R. K. & Cornwall, M. C. Role of noncovalent binding of 11-cis-retinal to opsin in dark adaptation of rod and cone photoreceptors. *Neuron* **29**, 749–755 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Engelmann and C. Koch for technical assistance; Y. Li for help in the early stage of the project; Y. J. Kim for help with data collection; M. Sommer, C. Enekel and M. Heck for critically reading the manuscript; and N. Krauss for support and sharing his expertise in crystallography. We thank U. Müller and the scientific staff of the Protein Structure Factory and the Freie Universität Berlin at beamlines BL 14.1 and BL 14.2 at BESSY for continuous support of the project. This work was supported by the Deutsche Forschungsgemeinschaft Sfb449 (to O.P.E.), Sfb740 (to O.P.E. and K.P.H.), DFG-KOEF international cooperation ER 294/1-1 (to O.P.E.) and F01-2004-000-10054-0 (to H.-W.C.). H.-W.C. was supported by CBNU funds for overseas research 2006–2007.

Author Information The coordinates and structure factors have been deposited in the Protein Data Bank under accession number 3CAP. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to O.P.E. (oliver.ernst@charite.de), H.-W.C. (hwchoe@chonbuk.ac.kr) or K.P.H. (klaus.peter.hofmann@charite.de).

METHODS

Crystallization. Disc membranes containing opsin were prepared from bovine eyes as described previously⁴³ and resuspended in 20 mM BTP (pH 7.5), 130 mM NaCl, 1 mM MgCl₂ and 10% sucrose. For crystallization, opsin was solubilized essentially as described, using 1% β -D-octylglucopyranoside¹³. Solubilized opsin with A_{280}/A_{362} ratios greater than four was used without further purification in crystallization screens by the sparse matrix crystallization method⁴⁴. Promising conditions were systematically screened further by changing protein concentration, pH and the concentration of precipitation agents. Opsin crystals could be grown by hanging-drop vapour diffusion method at 277 K using 24-well Linbro plates. Each hanging drop was prepared on a siliconized coverslip by mixing equal volumes (2 μ l each) of solubilized opsin (5 mg ml⁻¹, $\epsilon_{280} = 81,200 \text{ M}^{-1} \text{ cm}^{-1}$; ref. 45) and reservoir solution. The reservoir solution contained 2.8–3.2 M ammonium sulphate in 0.1 M MES or 0.1 M sodium acetate buffer, pH 5.6. Colourless opsin crystals appeared within 2–3 days and grew further for 5 days. Fully grown crystals had dimensions of $0.1 \times 0.1 \times 0.2 \text{ mm}^3$. On addition of 11-*cis*-retinal to the crystal, a colour change to red was observed, which vanished after illumination with white light, indicating that rhodopsin could be regenerated.

Structure solution and refinement. X-ray data collection of opsin crystal was performed at 100 K using cryoprotectant consisting of 90% (v/v) reservoir solution and 10% (w/v) trehalose. Diffraction data were collected at synchrotron beamline BL 14.2 of the Protein Structure Factory and Freie Universität Berlin at BESSY with a MAR-165 CCD detector. The crystal-to-detector distance was fixed at 210 mm. The rotation increment for each frame was 0.5° with an exposure time of 6 s. All images were indexed, integrated and scaled using HKL2000 (ref. 46). Opsin crystals belong to space group $H3$ ($a = 242.92 \text{ \AA}$, $b = 242.92 \text{ \AA}$, $c = 110.42 \text{ \AA}$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$).

Initial phases for opsin were obtained by molecular replacement using several models of rhodopsin monomers (PDB accessions 1U19, 1GZM and 1HZX) without cytoplasmic and extracellular regions as initial search trials. Molecular replacement was achieved using CCP4 program PHASER⁴⁷ by first placing the seven-transmembrane bundle of rhodopsin monomer (rotation function (RFZ): $Z = 3.7$; translational function (TFZ): $Z = 9.6$ as defined in PHASER) followed by a second (RFZ: $Z = 3.9$; TFZ: $Z = 21.5$). Initial cycles of simulated annealing refinement by the program CNS⁴⁸ were carried out maintaining two-fold non-crystallographic symmetry (NCS) constraints. After rigid body and B-factor refinement in CNS, major helix movements were clearly visible in both σ_A -weighted $F_o - F_c$ maps, as well as simulated annealing omit electron density maps. In later stages, the strict constraints were loosened within the two subunits of the dimer. Restrained, individual B-factors were refined and the crystal structure was finalized by CCP4 program REFMAC5 and CCP4 (ref. 47). The final model has agreement factors R_{free} and R_{cryst} of 26.6% and 22.9%, respectively. Manual rebuilding of the opsin model and electron density interpretation were performed using the program COOT⁴⁹. The structure was validated with the programs PROCHECK⁵⁰ and WHAT_CHECK⁵¹. All crystal structure superpositions of α traces were performed using CCP4 program LSQKAB⁴⁷. All molecular graphics representations were created using PyMol⁵².

43. Sachs, K., Maretzki, D. & Hofmann, K. P. Assays for activation of opsin by all-*trans*-retinal. *Methods Enzymol.* **315**, 238–251 (2000).

44. Jancarik, J. & Kim, S.-H. Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallogr.* **24**, 409–411 (1991).

45. Surya, A., Foster, K. W. & Knox, B. E. Transducin activation by the bovine opsin apoprotein. *J. Biol. Chem.* **270**, 5024–5031 (1995).

46. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).

47. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).

48. Brunger, A. T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).

49. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

50. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).

51. Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* **381**, 272 (1996).

52. DeLano, W. L. The PyMOL Molecular Graphics System. <<http://www.pymol.org>> (2002).

LETTERS

Rotational breakup as the origin of small binary asteroids

Kevin J. Walsh^{1,2}, Derek C. Richardson² & Patrick Michel¹

Asteroids with satellites are observed throughout the Solar System, from subkilometre near-Earth asteroid pairs to systems of large and distant bodies in the Kuiper belt. The smallest and closest systems are found among the near-Earth and small inner main-belt asteroids, which typically have rapidly rotating primaries and close secondaries on circular orbits. About 15 per cent of near-Earth and main-belt asteroids with diameters under 10 km have satellites^{1,2}. The mechanism that forms such similar binaries in these two dynamically different populations was hitherto unclear. Here we show that these binaries are created by the slow spinup of a 'rubble pile' asteroid by means of the thermal YORP (Yarkovsky–O'Keefe–Radzievskii–Paddack) effect. We find that mass shed from the equator of a critically spinning body accretes into a satellite if the material is collisionally dissipative and the primary maintains a low equatorial elongation. The satellite forms mostly from material originating near the primary's surface and enters into a close, low-eccentricity orbit. The properties of binaries produced by our model match those currently observed in the small near-Earth and main-belt asteroid populations, including 1999 KW₄ (refs 3, 4).

The angular momentum content from the primary's rotation and the secondary's orbit among small binaries suggests that the satellites were formed by rotational disruption after the body was pushed beyond its critical spin limit^{2,5}. Tidal encounters can account for near-critical spin rates and are efficient at forming binaries from rubble piles; however, they are even more efficient at subsequently dissociating those binaries as a result of repeated planetary encounters^{6,7}. In the main belt, the catastrophic disruption of an asteroid can produce binary systems, but they do not match the observed properties of small binaries^{8,9}. Radar observations of binary near-Earth asteroid (NEA) 1999 KW₄ show that the primary is oblate with a pronounced equatorial belt, the effective gravity at the equator is directed inward but is nearly zero, and its equatorial elongation is nearly unity^{3,4}. Owing to the quality of the observations, and the diagnostic 'top-like' shape of the primary, this system is a key constraint for binary formation models. The small main-belt asteroid (SMBAs) binaries have properties nearly identical to those of the NEA binaries, and both have an estimated frequency of ~15% (ref. 2). This suggests a common formation mechanism, which has not been identified so far.

One mechanism that operates on both NEAs and MBAs that may lead to the observed binaries is YORP-induced spinup, which arises from reflection and/or absorption and re-radiation of sunlight by the surface of an irregularly shaped asteroid^{10,11}. This effect accounts for the increase in the rotation rate of NEAs 2000 PH₅ and 1862 Apollo^{12–14}. The timescale for YORP spin alteration depends on the size R of the body (increasing with R^2), the distance a from the Sun (increasing with a^2), the body's thermal properties, and the

body's shape and obliquity. The YORP spinup/spindown timescale for kilometre-sized NEAs and MBAs is estimated to be between a few tens of thousands and a few million years, depending on the shape and makeup of the asteroid^{10,15}. Because of a notable abundance of both fast and slow rotators among NEAs and SMBAs, this effect seems to act widely¹⁶. However, it has never been shown whether gradual spinup leads to mass loss that can form binaries, and whether, if so, those binaries are a close match to observations.

We performed numerical simulations of YORP spinup of a cohesionless body consisting of ~1,000 self-gravitating rigid spheres. Several lines of evidence suggest that most kilometre-sized objects are rubble piles or gravitational aggregates, which means they lack cohesion but are non-fluid¹⁷. One indicator of the response of such bodies to stress is the angle of friction (ϕ) of the material. We modelled different kinds of rubble pile, ranging from a fluid-like body ($\phi \approx 0^\circ$), to a more typical terrestrial material ($\phi \approx 40^\circ$, here referred to as the nominal case)¹⁸. The model rubble piles consisted of either monodisperse spheres, or a simple bimodal distribution (meaning two different sizes of particles). Numerical experiments show that monodisperse rubble piles behave similarly to a body with $\phi \approx 40^\circ$, whereas ϕ for bimodal rubble piles depends on the relative particle sizes and their relative abundances within the body¹⁸. For our bimodal models, ϕ ranged from near 0° to $\sim 20^\circ$. We also tested another possible asteroid internal structure consisting of a rigid core of large particles surrounded by loose smaller particles. The representation of such a case by an angle of friction is not straightforward.

For the nominal case of $\phi \approx 40^\circ$, experiments were run with two different initial asteroid shapes: spherical and prolate. The prolate body had axis ratios of 2:1:1, and both shapes had initial spin periods of 4.4 h, longer than their stability limits for the body bulk density of 2.2 g cm^{-3} (where each particle had a density of 3.4 g cm^{-3}), so there was no immediate reshaping or collapsing. As the spin rate was increased (see below) and approached the critical spin limit, the spherical bodies became oblate, with mass moving from the poles to the equator. After this initial global reshaping of the spherical body, an equatorial belt of material remained, and subsequent mass loss originated from this region (Figs 1 and 2).

The fate of the ejected mass depends on the primary shape and the coefficient of restitution (the ratio of rebound to impact speed owing to energy dissipation when particles collide). In simulations with initially prolate bodies, the ejected mass does not readily accumulate into a satellite, because the mass that is ejected is lifted into a very shallow orbit barely above the surface of the primary and is easily disturbed by equatorial asymmetries in the prolate primary. In contrast, particles dislodged from spherical or oblate primaries quickly and efficiently accumulate into a satellite. For the most ideal parameters, namely $\phi \approx 40^\circ$ and a very low coefficient of restitution, the satellites accrete more than 90% of all ejected particles. In cases in which the primary

¹UMR 6202 Cassiopée, University of Nice-Sophia Antipolis, CNRS, Observatoire de la Côte d'Azur, BP 4229, 06304 Nice Cedex 4, France. ²Department of Astronomy, University of Maryland, College Park, Maryland 20742-2421, USA.

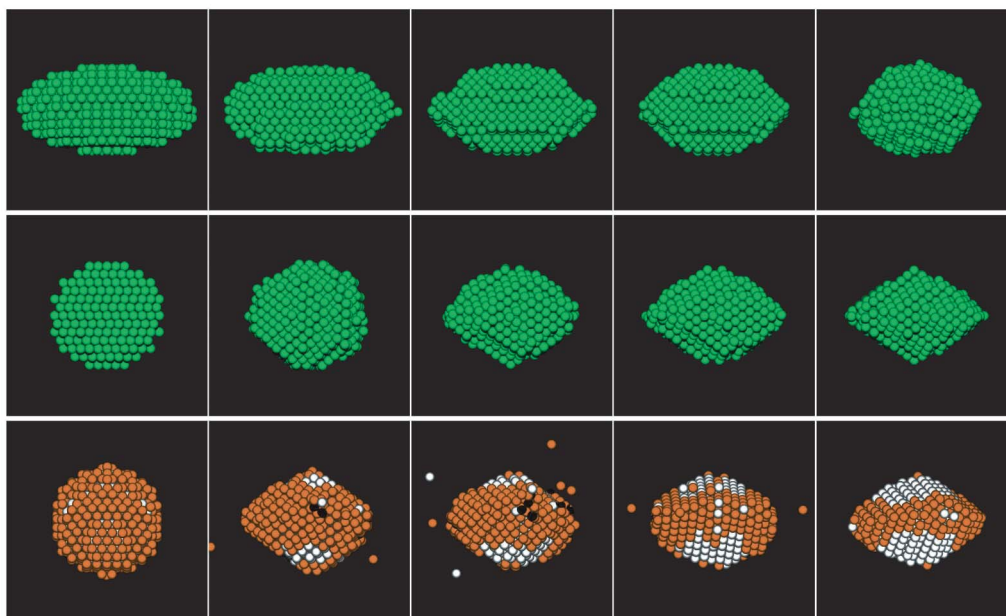


Figure 1 | Asteroid shape change during mass loss. The snapshots show the gradual change in shape that an initially prolate (top) and spherical (middle) body undergoes, as seen looking in the plane of the asteroid's equator. Also shown (bottom) is the movement and loss of a body's original surface particles (orange) and the exposure of the interior particles (white) during the binary formation. The top two rows show only the largest body in the simulation; ejected mass is not shown. Material that accumulates into a satellite does so slowly and from material lost from the equator of the

primary; there is no large-scale 'fission' event. The time between images is roughly 1,000 asteroid rotations for the top two rows, although for computational efficiency the simulations are sped up in comparison with the actual YORP effect. Prolate bodies become less elongated as particles are ejected from the ends of the long axis, reducing the critical rate for mass loss. Eventually prolate bodies become oblate, ending up with similar axis ratios to those of an initially spherical body.

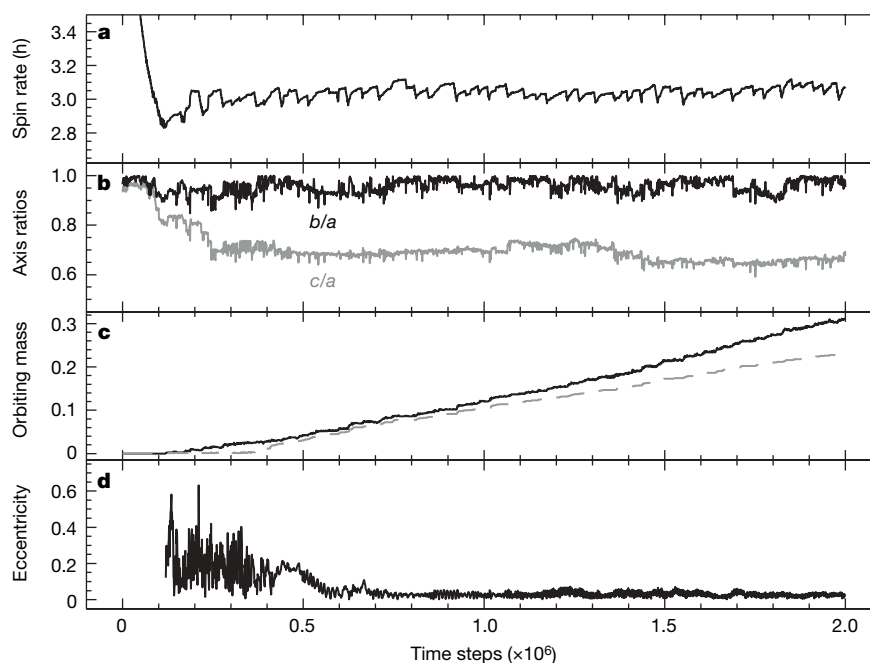


Figure 2 | Primary and secondary properties during satellite formation. **a**, Evolution of primary spin rate. **b**, Primary axis ratios (black, intermediate to long axis; grey, short to long axis). **c**, Mass loss (solid line) and satellite size (grey dashed line) as a fraction of progenitor mass. **d**, Satellite eccentricity. Plots are shown as a function of time steps of ~ 50 s. The originally spherical body becomes oblate after the increasing spin rate causes some mass loss. The newly oblate primary begins to accumulate mass in one satellite (dashed line in **c**), and the eccentricity quickly decreases to very low values. Initially prolate bodies show similar mass loss but do not accrete a satellite until

becoming oblate. The slow YORP spinup is modelled by applying small, discrete increases to the angular momentum of each particle making up the body, relative to the body's centre of mass. If any mass has been ejected or is in orbit, it is exempt from the angular momentum addition. The spin boosts are applied approximately every five rotation periods (for periods of ~ 3 h), allowing time for the body to equilibrate before more angular momentum is added to the system. If mass is lost between spin boosts, the next spin boost is delayed for at least about ten rotations, although these results were unchanged over a wide range of delay times between spin boosts.

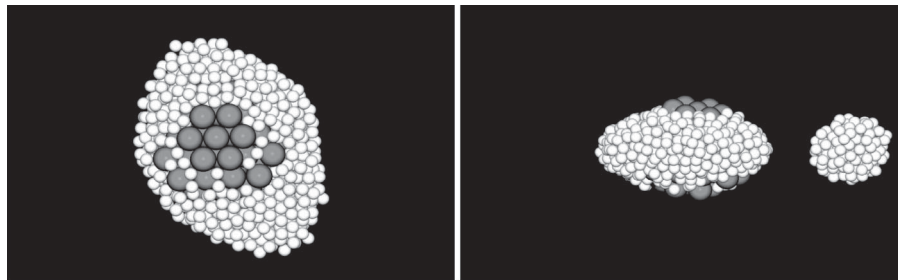


Figure 3 | Binary formation for an asteroid with a rigid core. Snapshot of binary formation for a body with a core of organized large particles (grey), making up $\sim 30\%$ of the total mass, surrounded by smaller particles (white). Shown are two views of the system at the same point in time: looking down the primary spin axis (left; only the primary is shown), and looking along the

plane of the primary's equator and the secondary's orbit (right). The core minimizes equatorial elongation growth, allowing satellite formation. In tests with a smaller core the body becomes very elongated and satellite formation is entirely frustrated.

shape is not initially spherical or oblate, satellite accumulation is delayed until the primary achieves a favourable shape.

The tendency of a gravitational aggregate to adopt an oblate shape as the angular momentum is increased is contrary to the evolution of fluid shapes (the classical Jacobi and Maclaurin figures), which become roughly prolate at rapid rotation rates. Simulations with $\phi \approx 20^\circ$ or $\phi \approx 0^\circ$ behaved most like the classical fluid case. The cases with $\phi \approx 0^\circ$ immediately adopted elongated shapes and maintained prolate shapes during mass loss, frustrating satellite formation for all test parameters. The intermediate test case, with $\phi \approx 20^\circ$, represented a transition, in which binary formation was possible but not very efficient. In our other test case of a substantial rigid core surrounded by smaller, loose particles, the core limited the overall elongation caused by motions of surface material arising from rapid rotation. Thus, a low equatorial elongation was maintained, permitting satellite formation (Fig. 3). Essentially, the minimum requirement for satellite formation is a low equatorial elongation, which was achieved in our models for aggregates with large non-zero angle of friction (which restricts reshaping), or aggregates with a substantial rigid core. In fact, Itokawa, the first asteroid in this size range to be visited by spacecraft, has a morphology suggestive of a large core surrounded by smaller debris¹⁹.

For our nominal case, massive satellites of minimum radius 0.2 primary radii (R_{pri}) formed in all simulations for which the lowest tested value of coefficient of restitution was used (0.2, where 1 is perfectly elastic and 0 is completely dissipative). Efficiency of satellite formation declined as the coefficient of restitution was increased, until, above a value of 0.6, no satellites formed. Evidently satellite accumulation is sensitive to energy dissipation during collisions, suggesting that collisions on the order of $0.2\text{--}0.5\text{ m s}^{-1}$ between asteroidal material dissipate significant amounts of energy. The actual value of the coefficient of restitution during collisions is not well constrained experimentally, but small-scale experiments suggest that it depends on the impact speed and material properties^{20,21}. Values as low as 0.2 can be expected, in particular for bodies with a certain degree of porosity²⁰, such as low-density asteroids and asteroids belonging to dark taxonomic type. Moreover, because the YORP timescale is inversely proportional to density, this model of binary formation is favoured for bodies with low bulk densities or for those consisting of collisionally dissipative materials. However, the YORP timescales are very short compared with dynamical lifetimes, so this mechanism may be indistinguishable between taxonomies. Currently all major taxonomic types are found among the observed binary systems, with no identifiable trends yet.

The exact properties of the secondary and its orbit depend strongly on when the YORP effect ceases to increase the spin of the primary and send mass to the secondary¹⁸. In our simulations, when secondaries grow to $0.3R_{\text{pri}}$, the orbital semimajor axes are between $2.0R_{\text{pri}}$ and $4.5R_{\text{pri}}$, eccentricities are all below 0.15, and the equatorial elongations of the primaries are all below 1.2. Most (70–90%) of the particles

forming the secondary originate from the surface of the primary. After the secondary has formed, 15–35% of the primary's surface is material that originated below the surface and is exposed mostly near the poles of the primary, whereas the equator is still largely covered with original surface material (see Fig. 1). The near absence of observed binary systems with very large secondaries, larger than about half the size of the primary, suggests that mass transfer stops at some point. Our simulations only model the gradual spinup of a single asteroid, and not the additional complex effects that a large secondary in a close orbit may produce, such as the binary YORP effect (BYORP, a radiation effect operating on the system rather than on the primary only), tidal interactions or continued reshaping of the primary. The long-term fate of the system therefore depends on the evolution of the binary, with the BYORP effect or planetary tides possibly splitting the system, leaving behind a rapidly rotating primary¹⁸.

The observed NEA and SMBA binary fraction ($\sim 15\%$) is probably a balance between YORP spinup and known or suspected dynamical sinks (planetary tides and BYORP). The similarities between binaries in the dynamically distinct NEA and SMBA populations arise from their shared minimum physical requirements for binary formation by means of YORP, properties that must be found among a larger population of asteroids that participate in the binary formation–destruction cycle. The requirements include a non-zero angle of friction for the component material (or a rigid core that resists reshaping under stress from rapid rotation), allowing oblate or spherical shapes to be maintained near the critical spin limit, and subsequently permitting the formation of stable satellites (which itself is dependent on a certain degree of collisional dissipation in the component material). These systems may be particularly attractive targets for space missions, because of the exposure of some fresh surface by the movement and removal of surface material from the poles to the equator of the asteroid.

Received 14 March; accepted 29 April 2008.

1. Pravec, P. *et al.* Photometric survey of binary near-Earth asteroids. *Icarus* **181**, 63–93 (2006).
2. Pravec, P. & Harris, A. W. Binary asteroid population. *Icarus* **190**, 250–259 (2007).
3. Ostro, S. J. *et al.* Radar imaging of binary near-Earth asteroid (66391) 1999 KW4. *Science* **314**, 1276–1280 (2006).
4. Scheeres, D. J. *et al.* Dynamical configuration of binary near-Earth asteroid (66391) 1999 KW4. *Science* **314**, 1280–1283 (2006).
5. Richardson, D. C. & Walsh, K. J. Binary minor planets. *Annu. Rev. Earth Planet. Sci.* **34**, 47–81 (2006).
6. Walsh, K. J. & Richardson, D. C. Binary near-Earth asteroid formation: Rubble pile model of tidal disruptions. *Icarus* **180**, 201–216 (2006).
7. Walsh, K. J. & Richardson, D. C. A steady-state model of NEA binaries formed by tidal disruption of gravitational aggregates. *Icarus* **193**, 553–566 (2008).
8. Michel, P. *et al.* Collisions and gravitational reaccumulation: forming asteroid families and satellites. *Science* **294**, 1696–1700 (2001).
9. Durda, D. D. *et al.* The formation of asteroid satellites in catastrophic impacts: Results from numerical simulations. *Icarus* **167**, 382–396 (2004).
10. Rubincam, D. P. Radiative spin-up and spin-down of small asteroids. *Icarus* **148**, 2–11 (2000).

11. Paddack, S. J. & Rhee, J. W. Rotational bursting of interplanetary dust particles. *Geophys. Res. Lett.* **2**, 365–367 (1975).
12. Lowry, S. C. *et al.* Direct detection of the asteroidal YORP effect. *Science* **316**, 272–274 (2007).
13. Taylor, P. A. *et al.* Spin rate of asteroid (54509) 2000 PH5 increasing due to the YORP effect. *Science* **316**, 274–277 (2007).
14. Kaasalainen, M., Durech, J., Warner, B. D., Kugly, Y. N. & Gaftonyuk, N. N. Acceleration of the rotation of asteroid 1862 Apollo by radiation torques. *Nature* **446**, 420–422 (2007).
15. Čuk, M. Formation and destruction of small binary asteroids. *Astrophys. J.* **659**, 57–60 (2007).
16. Pravec, P. & Harris, A. W. Fast and slow rotation of asteroids. *Icarus* **148**, 12–20 (2000).
17. Richardson, D. C., Leinhardt, Z. M., Melosh, H. J., Bottke, W. F. & Asphaug, E. in *Asteroids III* (eds Bottke, W. F. Jr, Cellino, A., Paolicchi, P. & Binzel, R. P.) 501–515 (Univ. of Arizona Press, Tucson, AZ, 2002).
18. Richardson, D. C., Elankumaran, R. E. & Sanderson, R. E. Numerical experiments with rubble piles: equilibrium shapes and spins. *Icarus* **173**, 349–361 (2005).
19. Fujiwara, A. *et al.* The rubble-pile asteroid Itokawa as observed by Hayabusa. *Science* **312**, 1330–1334 (2006).
20. Supulver, K. D., Bridges, F. G. & Lin, D. N. C. The coefficient of restitution of ice particles in glancing collisions: Experimental results for unfrosted surfaces. *Icarus* **113**, 188–199 (1995).
21. Fujii, Y. & Nakamura, A. M. Compaction and fragmentation of porous targets at low velocity collisions. *Lunar Planet. Sci. Conf. XXXVIII*, abstract 1525 (2007).

Acknowledgements We thank W. F. Bottke and A. Harris for their constructive reviews. K.J.W. and D.C.R. acknowledge support from the National Science Foundation under grants AST0307549 and AST0708110. K.J.W. and P.M. also had the support of the European Space Agency's Advanced Concepts Team on the basis of the Ariadna study 07/4111 'Asteroid Centrifugal Fragmentation', and of the French Programme National de Planétologie. K.W. is also supported by the Henri Poincaré fellowship at the Observatoire de la Côte d'Azur, Nice, France. We acknowledge the use of the Mésocentre de Calcul-SIGAMM hosted at the Observatoire de la Côte d'Azur, Nice, France. Some simulations were performed at the University of Maryland with the Department of Astronomy borg cluster and the Office of Information Technology High Performance Computing Cluster. Raytracing for Figs 1 and 3 was performed with the Persistence of Vision Raytracer.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.W. (kwals@oca.eu).

LETTERS

Volatile content of lunar volcanic glasses and the presence of water in the Moon's interior

Alberto E. Saal¹, Erik H. Hauri², Mauro Lo Cascio¹, James A. Van Orman³, Malcolm C. Rutherford¹ & Reid F. Cooper¹

The Moon is generally thought to have formed and evolved through a single or a series of catastrophic heating events¹, during which most of the highly volatile elements were lost. Hydrogen, being the lightest element, is believed to have been completely lost during this period². Here we make use of considerable advances in secondary ion mass spectrometry³ to obtain improved limits on the indigenous volatile (CO₂, H₂O, F, S and Cl) contents of the most primitive basalts in the Moon—the lunar volcanic glasses. Although the pre-eruptive water content of the lunar volcanic glasses cannot be precisely constrained, numerical modelling of diffusive degassing of the very-low-Ti glasses provides a best estimate of 745 p.p.m. water, with a minimum of 260 p.p.m. at the 95 per cent confidence level. Our results indicate that, contrary to prevailing ideas, the bulk Moon might not be entirely depleted in highly volatile elements, including water. Thus, the presence of water must be considered in models constraining the Moon's formation and its thermal and chemical evolution.

Volatile elements provide insight into models of planet formation, and play a fundamental role in planetary evolution through their influence on melting⁴, viscosity⁵, magma crystallization⁶ and volcanic eruption. An important conclusion resulting from the Apollo and Luna programmes is that the Moon is deficient in highly volatile elements relative to the Earth. This is especially the case for hydrogen, which is thought to have been completely lost during the giant collision event that generated the Moon².

The volatile budget of the lunar mantle can, at present, only be reconstructed from the record preserved in the mare basalts and the lunar volcanic glasses, the most primitive basalts from the Moon. Reconstructing the volatile content of the lunar mantle from basaltic melts is compromised by volatile degassing at the time the lava erupted and subsequent contamination from external sources such as low-pressure condensation following impact, solar wind implantation, and assimilation or sublimation of cometary or meteoritic material⁷. Over the past 40 years there have been considerable efforts to measure the volatile contents in these lunar samples^{8–15} (see Supplementary Information). From these studies, we have reliable evidence that there is an indigenous component containing sulphur, and, to a lesser extent, chlorine, fluorine and carbon in these materials. Yet the evidence for indigenous H₂O in the lunar samples has remained elusive, consistent with the general consensus that the Moon is anhydrous.

Two main factors have limited the study of volatile abundances in lunar samples: first, most of the existing data represent bulk sample analyses, which makes it difficult to determine whether the volatile elements measured were indigenous to the glass beads or foreign (implanted, condensed or added during alteration of the sample); and second, the *in situ* analytical techniques used (Fourier transform infrared (FTIR), electron microprobe and secondary ion mass

spectrometry (SIMS)) have had relatively high detection limits, particularly for H₂O and CO₂. Recent substantial advances in SIMS provide improved detection limits for H₂O, CO₂, F, S and Cl, up to two orders of magnitude lower than electron microprobe, FTIR and earlier SIMS instrumentations (see Supplementary Information). Here we report, by virtue of the new SIMS technique³, improved limits on the indigenous volatile (CO₂, H₂O, F, S, Cl) contents of the lunar volcanic glasses and evaluate the processes controlling their variation within and between glass beads. Our results represent the first evidence for the presence of indigenous water in the lunar interior.

We investigated three main compositional groups of glasses: very-low-Ti and low-Ti glasses (sample 15427,41), and high-Ti glasses (sample 74220,864). The glassy spherules range in size from 100 to 300 µm for 74220,864 and from 200 to 400 µm with one outlier at 700 µm for 15427,41. Not all of the glass beads are completely glassy; some of them, usually the larger beads, show crystallization of either olivine (15427,41) or olivine and ilmenite (74220,864). Also, we were able to recognize four of the five compositional subgroups (A, B, C and D; group E was not sampled) of Delano's very-low-Ti glasses¹⁶. The major and trace element contents of the lunar volcanic glasses are consistent with previously reported data and indicate that the glasses analysed in this study represent volcanic rather than impact glasses^{14,17,18} (see Supplementary Tables 1 and 2, Fig. 1 and the Methods section for detailed discussion on the analytical methods).

Essentially all volcanic glasses have carbon content (reported as CO₂) within 2σ (standard deviation) of the detection limit, with concentrations of $\sim 6 \pm 7$ p.p.m. CO₂ after background correction. Carbon will not be considered further, beyond mentioning that our values give an upper limit on the concentration of carbon dissolved in the volcanic glasses. However, it is important to point out that two high-Ti glasses have 13 ± 7 and 18 ± 7 p.p.m. CO₂; if confirmed, these would be the first direct evidence for measurable dissolved carbon in any of the lunar volcanic glasses. The other volatiles, after background correction, have measurable abundances of H₂O (4–46 p.p.m.), F (4–40 p.p.m.), S (115–576 p.p.m.) and Cl (0.06–2 p.p.m.) (see Supplementary Table 1).

There are no clear correlations between the volatile contents and major and trace element contents when we consider all the compositional groups (very-low-, low- and high-Ti glasses). Yet each group has a specific range in F and S contents, suggesting differences in their initial volatile content that have not been completely erased by the degassing process acting during melt transport and eruption.

An important feature of the data is the clear correlation between H₂O, other volatile species and major elements among the very-low-Ti glasses (Fig. 1; Supplementary Figs 2, 3 and 4). These correlations indicate that H₂O in the glasses is indigenous, not a product of solar wind implantation or laboratory contamination, and support the

¹Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA. ²Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington DC 20015, USA. ³Department of Geological Sciences Case Western Reserve University, Cleveland, Ohio 44106, USA.

hypothesis that there were significant differences in the initial volatile content and/or the extent of degassing among the lunar glasses. The correlations also suggest that the major element composition of the melt may have influenced the degassing dynamics through volatile

diffusion and bubble formation kinetics (see Supplementary Figs 5 and 6, and detailed discussion in the Supplementary Information on the implications of our results).

To evaluate further the hypothesis that volatiles in the lunar volcanic glasses are indigenous but affected by partial degassing during eruption, we measured radial concentration profiles for the volatiles within a single very-low-Ti glass bead (see Fig. 2; Supplementary Table 4 and the Methods section for detailed discussion on the analytical methods). Volatile contents in this bead decrease systematically from core to rim. This decrease is especially significant for H_2O , which has a concentration of ~ 30 p.p.m. in the centre, decreasing steadily to ~ 14 p.p.m. near the rim. Similar depletion towards the rim of the glass bead is observed in the data for F, Cl and S. This observation supports the hypothesis that H_2O and the other volatiles were indigenous to the glass bead but were affected by degassing during eruption. The depletion within the single glass bead reproduces essentially the same correlations between volatile elements as those defined by the volatile contents of all the very-low-Ti glasses (see Supplementary Fig. 4). This reinforces the hypothesis that degassing has been the main process affecting the primitive volatile contents of the lunar volcanic glasses.

To evaluate diffusive volatile loss from the lunar volcanic glasses during eruption, we applied a model of diffusion from a sphere with concomitant surface evaporation to the measured volatile concentration profiles within the selected glass bead¹⁰ (Fig. 2). We emphasize that our calculations consider only volatile loss by diffusive degassing from the time of eruption until deposition, and do not consider volatile loss during the period of bubble formation and growth preceding magma fragmentation and eruption. Furthermore, we assume that there was no significant change in the distribution of volatiles within the glass beads during the period of 3.4–3.8 Gyr following their deposition¹⁰. This assumption is reasonable because at the average temperature on the sunlit side of the Moon (-20 to 0°C) diffusion in the silicate glasses is negligible, and the beads have been exposed directly to the lunar surface for no more than 30 Myr (refs 19, 20). The crucial input parameters of the diffusion model are the sphere radius, the temperature of the melt during eruption, the cooling rate, the initial volatile concentrations in the glass beads, the diffusion coefficients of the volatiles within the melt, and the rates of evaporation at the surface of the melt sphere. We examined a wide range of cooling rates and times, and considered the concentration profiles for all volatiles simultaneously to determine the parameters that provided the best fit (see Supplementary Information for details on the model).

The degassing model provides excellent fits to the measured radial concentration profiles for all species simultaneously, at reasonable cooling rates and cooling times (Fig. 2). We found that cooling rates greater than 4 K s^{-1} did not provide acceptable fits for all volatile species (see Supplementary Fig. 7). Cooling rates of 2 to 3 K s^{-1} over a period of ~ 2 to 5 min between eruption and quenching provided the best fit to all volatile profiles together; at these conditions the glass bead loses approximately 19% S, 45% F, 57% Cl and 98% H_2O . The initial H_2O content is not precisely constrained by the modelling, but the best fit, again considering cooling rates and times that allow acceptable fits for all volatile profiles simultaneously, is obtained for an initial H_2O content of 745 p.p.m., and the H_2O content must be at least 260 p.p.m. to obtain an acceptable fit (at the 95% confidence level).

Several different origins of the lunar volatiles might be proposed. An undifferentiated lunar mantle reservoir comparatively rich in volatiles might exist at a depth greater than that of the postulated lunar magma ocean, or a less-processed cumulate that kept the primitive lunar mantle signature may have retained significant volatile elements¹⁸. Another possibility would be the recycling of an ilmenite-rich dense layer enriched in incompatible and volatile elements formed during the late-stage cooling and crystallization of the lunar magma ocean²¹. This process may have been responsible for the

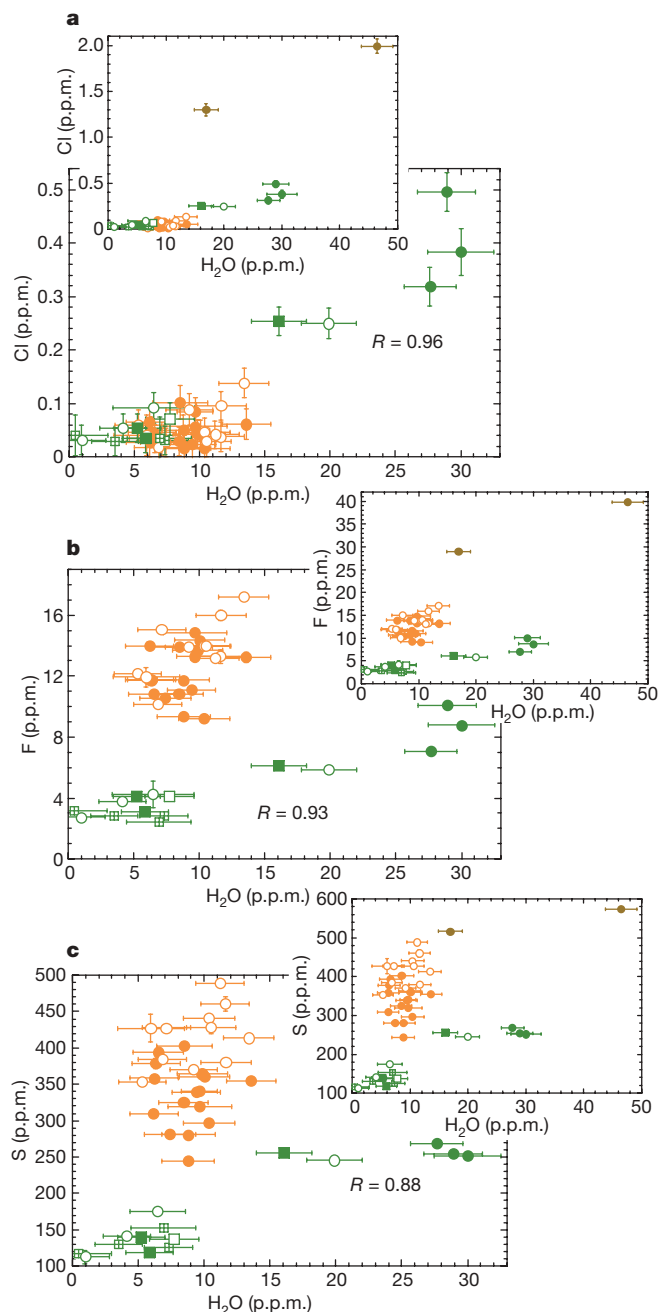


Figure 1 | Correlations between water and other volatile content in the lunar volcanic glasses. **a**, Chlorine; **b**, fluorine; and **c**, sulphur against H_2O content. Green symbols represent the very-low-Ti glasses: group B and C (filled circles), A (filled squares) and D (crossed open squares); partially crystallized glasses in each group are represented by open circles (group B and C) and open squares (group A). Orange filled circles represent high-Ti glasses; open orange circles indicate partially crystallized glasses. Inset shows all the glasses including the two low-Ti glasses (filled brown circles). Error bars represent standard deviation (2σ) uncertainties. Volatile contents are reported in parts per million. There are significant correlations between the volatile contents measured for the very-low-Ti glasses (see Supplementary Tables 1 and 3 and Fig. 2). Note that although the high-Ti glasses have similar Cl and H_2O contents to the very-low-Ti glasses, they have significantly higher F and S contents, suggesting different initial volatile concentrations between the different major compositional groups of glasses.

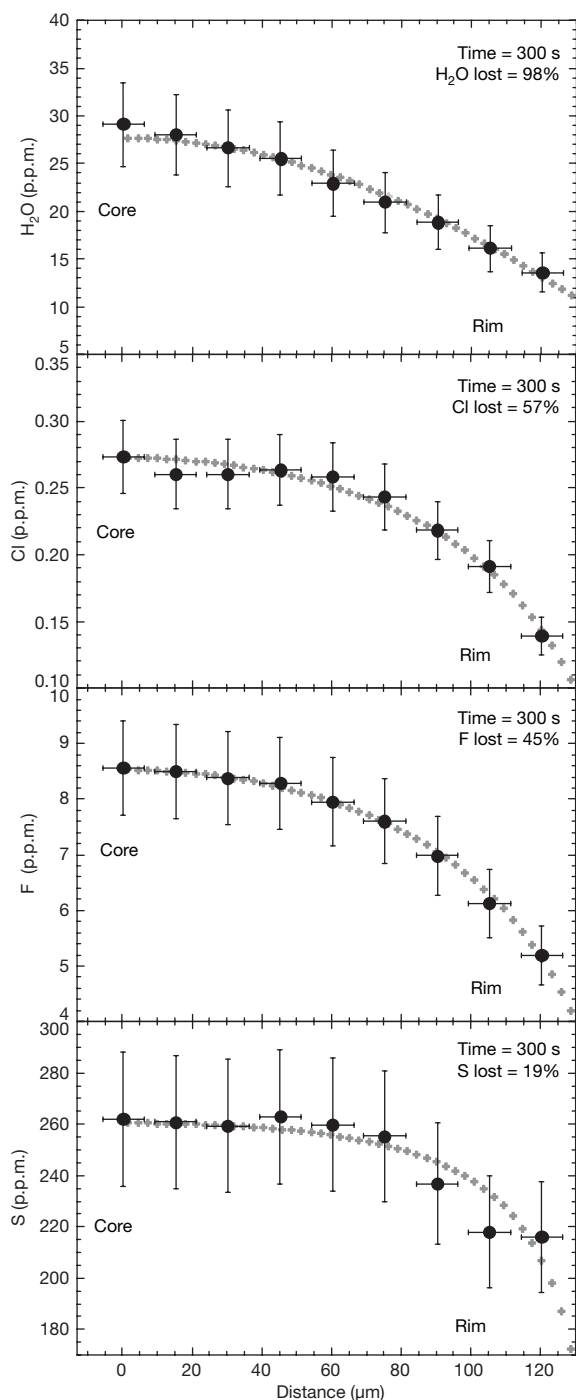


Figure 2 | Volatile concentration profiles from core to rim in a single bead of the very-low-Ti glasses. The glass used is 'Green #5' (see Supplementary Tables 1 and 4). The black filled circles represent the measured profiles, the grey filled crosses define the output data of the model describing the diffusive volatile loss from a homogeneous sphere with concomitant surface evaporation; see text for details and Supplementary Tables 4 and 5 (Case 1) for model parameters. Error bars represent standard deviation (2σ) uncertainties, and the size of the rastered ion beam. Volatile contents are in parts per million. To avoid any possible influence of ions implanted by solar wind, we did not analyse the outermost rim of the glass bead; the measured concentration profiles stop 18 μm before reaching the rim of the bead. The sharp decrease in volatile content from core to rim for H_2O , Cl, F and S suggests that the volatile contents are indigenous to the Moon and were affected by degassing during magma eruption. As expected, H_2O has the largest observed variation from core to rim, followed by Cl, F and S, indicating progressively less degassing for those elements, respectively.

re-fertilization of volatile elements into the source of the volcanic glasses.

Previous hypotheses suggest that the volatile elements either survived or were accreted after the giant impact event that probably led to the formation of the Moon. Pahlevan and Stevenson²² suggested that the proto-Earth (having significant amounts of water) and proto-lunar disk would have diffusively equilibrated after the giant impact, and that the volatile depletion of the Moon may be explained by hydrodynamic escape from the proto-lunar disk driven by an outflow of hydrogen previously accreted to the Earth. Our results suggest either that hydrodynamic escape was not complete or that a significant amount of water was accreted to the Earth–Moon system immediately after the giant impact^{23,24}. The existence of zircons on the Earth that are as old as 4,325 Myr and have oxygen isotopes of 6.5‰ provides evidence for the presence of liquid water near the Earth's surface within ~ 230 Myr of the Earth's accretion^{25–27}. This observation strongly suggests that either the Earth–Moon system retained significant amounts of water after the giant impact, or that volatile-rich material accreted to both the Earth and Moon within a narrow time window after the giant impact but before 4.3 Gyr. At this point we do not have enough information to support or discard either of these hypotheses. Our results suggest that, contrary to the prevailing ideas, the bulk Moon might not be entirely depleted in highly volatile elements, and the presence of volatiles, especially water, must be included in models constraining the Moon's formation and its thermal and chemical evolution.

METHODS SUMMARY

Modelling. To evaluate diffusive volatile loss from the lunar volcanic glasses during eruption, we applied a model of diffusion from a sphere using a temperature-dependent diffusion coefficient with concomitant surface evaporation^{10,28}. The equations and boundary conditions were solved numerically using a forward-time, centred-space finite-difference scheme²⁸ for each element independently. We selected the model outputs that minimized the sum of χ^2 for all of the volatiles together, using the optimum values of initial concentration and evaporation constant determined for each volatile at the cooling rate and cooling time of interest (see Supplementary Information).

Analytical method. The glass beads were individually selected and mounted in indium and analysed for major elements using the Cameca SX100 electron microprobe, and for trace and volatile elements using the Cameca IMS 6f and the NanoSIMS 50L (see online Methods and Supplementary Tables 1, 2 and 4)^{3,29}. We measured the volatile contents by SIMS using methods recently developed for the simultaneous microanalysis of trace amounts of H_2O , CO_2 , F, S and Cl in glasses and nominally anhydrous minerals³.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 February; accepted 28 April 2008.

- Canup, R. M. Dynamics of lunar formation. *Annu. Rev. Astron. Astrophys.* **42**, 441–475 (2004).
- Lucey, P. et al. in *New Views of the Moon. Reviews in Mineralogy and Geochemistry* Vol. 60 (eds Jolliff, B. L., Wieczorek, M. A., Shearer, C. K. & Niel, C. R.) 83–219 (Mineralogical Society of America, Chantilly, Virginia, 2006).
- Hauri, E. H., Gaetani, G. A. & Green, T. H. Partitioning of water during melting of the Earth's upper mantle at H_2O -undersaturated conditions. *Earth Planet. Sci. Lett.* **248**, 715–734 (2006).
- Gaetani, G. A. & Grove, T. L. The influence of water on melting of mantle peridotite. *Contrib. Mineral. Petrol.* **131**, 323–346 (1998).
- Hirth, G. & Kohlstedt, D. Water in the oceanic upper mantle: implications for rheology, melt extraction and the evolution of the lithosphere. *Earth Planet. Sci. Lett.* **144**, 93–108 (1996).
- Asimow, P. D. & Langmuir, C. H. The importance of water to oceanic mantle melting regimes. *Nature* **421**, 815–820 (2003).
- Wieczorek, M. A. et al. in *New Views of the Moon. Reviews in Mineralogy and Geochemistry* Vol. 60 (eds Jolliff, B. L., Wieczorek, M. A., Shearer, C. K. & Niel, C. R.) 221–364 (Mineralogical Society of America, Chantilly, Virginia, 2006).
- Epstein, S. & Taylor, H. P. The isotopic composition and concentration of water, hydrogen and carbon in some Apollo 15 and 16 soils and in the Apollo 17 orange soil. *Geochim. Cosmochim. Acta* **2**, 1559–1575 (1973).
- Gibson, E. K. & Moore, G. W. Volatile-rich lunar soil: evidence of possible cometary impact. *Science* **179**, 69–71 (1973).
- Fogel, R. A. & Rutherford, M. J. Magmatic volatiles in primitive lunar glasses; I, FTIR and EPMA analyses of Apollo 15 green and yellow glasses and revision of the

- volatile-assisted fire-fountain theory. *Geochim. Cosmochim. Acta* **59**, 201–215 (1995).
11. Goldberg, R. H., Burnett, D. S. & Tombrello, T. A. in *Proc. 6th Lunar Planet. Sci. Conf.* Vol. 2 2189–2200 (Pergamon, New York, 1975).
 12. Jovanovic, S. & Reed, G. W. Jr. in *Proc. 6th Lunar Planet. Sci. Conf.* Vol. 2 1737–1751 (Pergamon, New York, 1975).
 13. Gibson, E. K. Jr, Brett, R. & Andrawes, F. in *Proc. 8th Lunar Planet. Sci. Conf.* Vol. 2 1417–1428 (Pergamon, New York, 1977).
 14. Delano, J. W., Hanson, B. Z. & Watson, E. B. in *Proc. 25th Lunar Planet. Sci. Conf.* 325–326 (Lunar and Planetary Institute, Houston, 1994).
 15. Elkins-Tanton, L. T., Chatterjee, N. & Grove, T. L. Magmatic processes that produced lunar fire fountains. *Geophys. Res. Lett.* **30**, 20–21 (2003b).
 16. Delano, J. W. in *Proc. 10th Lunar Planet. Sci. Conf.* Vol. 1 275–300 (Pergamon, New York, 1979).
 17. Delano, J. W. Pristine lunar glasses; criteria, data, and implications. *J. Geophys. Res.* **91**, D201–D213 (1986).
 18. Shearer, C. K. et al. in *New Views of the Moon. Reviews in Mineralogy and Geochemistry* Vol. 60 (eds Jolliff, B. L., Wieczorek, M. A., Shearer, C. K. & Niel, C. R.) 365–518 (Mineralogical Society of America, Chantilly, Virginia, 2006).
 19. Lakatos, S., Heymann, D. & Yaniv, A. Green spherules from Apollo 15: Inferences about their origin from inert gas measurements. *The Moon* **7**, 132–148 (1973).
 20. Eugster, O. et al. in *Proc. 8th Lunar Planet. Sci. Conf.* Vol. 2 3059–3082 (Pergamon, New York, 1977).
 21. Hess, P. C. & Parmentier, E. M. A model for the thermal and chemical evolution of the Moon's interior; implications for the onset of mare volcanism. *Earth Planet. Sci. Lett.* **134**, 501–514 (1995).
 22. Pahlevan, K. & Stevenson, D. J. Equilibration in the aftermath of the lunar-forming giant impact. *Earth Planet. Sci. Lett.* **262**, 438–449 (2007).
 23. Bottke, W. F., Levison, H. F., Nesvorn, D. & Dones, L. Can planetesimals leftover from terrestrial planet formation produce the lunar Late Heavy Bombardment? *Icarus* **190**, 203–223 (2007).
 24. Ryder, G. Mass flux in the ancient Earth–Moon system and benign implications for the origin of life on Earth. *J. Geophys. Res.* **107**, 5022, doi:10.1029/2001JE001583 (2002).
 25. Cavosie, A. J., Valley, J. W., Wilde, S. A. & EIMF. Magmatic $\delta^{18}\text{O}$ in 4400–3900 Ma detrital zircons: A record of the alteration and recycling of crust in the Early Archean. *Earth Planet. Sci. Lett.* **235**, 663–681 (2005).
 26. Wilde, S. A., Valley, J. W., Peck, W. H. & Graham, C. M. Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. *Nature* **409**, 175–178 (2001).
 27. Mojzsis, S. J., Harrison, T. M. & Pidgeon, R. T. Oxygen-isotope evidence from ancient zircons for liquid water at the Earth's surface 4,300 Myr ago. *Nature* **409**, 178–181 (2001).
 28. Crank, J. *The Mathematics of Diffusion* (Oxford Univ. Press, Oxford, UK, 1975).
 29. Shimizu, N. & Hart, S. R. Application of the ion probe to geochemistry and cosmochemistry. *Annu. Rev. Earth Planet. Sci.* **10**, 483–526 (1982).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Delano for guidance on sample selection, P. Hess for exchange of ideas, M. Chaussidon, J. Longhi and T. Grove for reviews, J. Wang and J. Devine for technical assistance, and the NASA Cosmochemistry programme and the NASA Astrobiology Institute for support.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.E.S. (asaal@brown.edu).

METHODS

The lunar glasses were mounted in indium and analysed for major elements using the Cameca SX100 electron microprobe at the Department of Geological Sciences, Brown University. We made the analyses using 15 kV accelerating voltage, 10 nA beam intensity, a 5–10 μm defocused beam and PAP correction procedures³⁰. Most major elements reported have precision (2σ) of 1–1.5%, with the exception of Na_2O 20–60%, K_2O 20–80%, Cr_2O_3 ~15%, TiO_2 5–15%, P_2O_5 <100%, MnO <20%, which represents the average of the standard deviation of three to eight replicated analyses on a single glass bead. We reported Cr_2O_3 , K_2O and P_2O_5 contents obtained using the Cameca 6f ion probe at DTM, Carnegie Institution of Washington. The precision for the trace elements measured by ion probe is better than 15% (2σ) and represents the average of the standard deviation of three to four replicated analyses on a single glass bead. SIMS calibrations are regressions of ion probe signals compared with known concentrations.

In this study, we plotted standard trace element (and volatile) concentrations against measured trace element (volatile)/ ^{30}Si ratios^{3,29}. We measured the volatile contents by SIMS using a Cameca IMS 6f and the NanoSIMS 50L at DTM, Carnegie Institution of Washington, using methods recently developed for the microanalysis of trace amounts of H_2O , CO_2 , F, S and Cl in glasses and nominally anhydrous minerals^{3,31–33}. For the Cameca IMS 6f a typical 10-min measurement for volatile abundances is made on a singly polished specimen using a Cs^+ primary beam (~14 nA accelerated to 10 kV) with collection of negatively charged secondary ions. We used a 10- μm primary beam rastered at 25 μm , which results in a 35- μm crater. We pre-sputter for 3–5 min before analysis. During this time, we monitored secondary ion images of ^{12}C , ^{17}OH , ^{19}F , ^{32}S and ^{35}Cl projected on the channel plate. This procedure helped to avoid inclusions and cracks, which appear as bright features on the projected image (especially the ^{12}C image), and dendritic crystals (dark shapes on ^{32}S image). After each beam spot had been carefully examined a field aperture was inserted to permit transmission of ions only from the central 10 μm of the 35- μm crater (that is, always 12.5 μm away from the crater edge), thus avoiding transmission of ions from the edge of the sputter crater and the surface of the sample. Counting times were 10 s for ^{12}C and 5 s for all other elements. Pressure in the ion probe sample chamber was $\sim 6 \times 10^{-10}$ torr during the analyses. The sample mount was placed in the ion probe sample chamber for ~12 h before analysis; this approach allowed the best detection limits for H_2O . Synthetic forsterite (Allied, <0.4 p.p.m. H_2O by FTIR) was used for the determination of H_2O detection limits^{3,31–33}. All the lunar glasses were contained on a single sample mount together with synthetic forsterite; the H_2O detection limit on this particular sample mount was 6.4 p.p.m., as determined by five separate measurements interspersed with analyses of the lunar glasses. These measurements of synthetic forsterite yielded an average ^{16}OH count rate of 150 counts per second with 3,750 total counts per analysis and a Poisson limit on the precision of 3.2% (2σ), which is much lower than the actual reproducibility of the detection limit (~25%, 2σ). As a result, none of the analyses is limited in any way by counting statistics; the limiting factor is reproducibility of

the detection limit. Similar statistics are obtained for the other volatile elements on forsterite; detection limits for F, S and Cl are ~0.09 p.p.m., ~0.27 p.p.m. and ~0.03 p.p.m. respectively.

Calibrations for H_2O and other volatiles were verified for glasses and nominally anhydrous mineral standards before each analytical session; abundances for the other volatile elements (CO_2 , F, S, Cl) were calculated from the calibrations performed on glass standards. The reported volatile concentrations of lunar glasses are obtained by simply subtracting the detection limit from the measured concentrations, and the uncertainties are calculated by propagating the errors in the detection limit and the counting statistics. After the measurement by the IMS 6f, we re-polished the sample and measured the core to rim volatile variation with the Cameca NanoSIMS 50L at DTM. We followed the same analytical method used on the Cameca 6f, with the exception that crater edge contamination was eliminated by electronic gating rather than a field aperture. A typical 15-min measurement used a Cs^+ primary beam (~3 nA accelerated to 8 kV) with collection of negatively charged secondary ions on multiple detectors. We used an 800-nm primary beam rastered at $12 \times 12 \mu\text{m}$ area divided into 64×64 pixels with 140- μs dwell time per pixel, and an electronic gating on the central $4.5 \times 4.5 \mu\text{m}$. We pre-sputter for 6 min before analysis and collected the data on six detectors in multi-collection mode: ^{12}C , ^{16}OH , ^{19}F , ^{30}Si , ^{32}S , ^{35}Cl at mass resolving power of ~6,000 (sufficient to resolve ^{16}OH from ^{17}O). Counting times were 1 s for all masses, and we collected 100 ratios (100 s total counting time). Pressure in the ion probe sample chamber was $\sim 3 \times 10^{-10}$ torr or less during the analyses. We used synthetic forsterite (Allied, <0.4 p.p.m. H_2O by FTIR) to establish the H_2O detection limit of 13 p.p.m., as determined by five separate measurements interspersed with analyses of the lunar glasses. Count rates for ^{16}OH were 540 counts s^{-1} , with 54,000 total counts per analysis and a Poisson precision limit of 0.002% (2σ), well below the uncertainty in the detection limits ($\pm 8\%$ 2σ). The limiting factor on the precision is the reproducibility of the detection limit. The reported volatile concentrations of lunar glasses are obtained by simply subtracting the detection limit from the measured concentrations. For the uncertainties, we assigned a conservative 15% (2σ) for H_2O and 10% (2σ) for F, S and Cl on all the measured concentration profiles, which represent the higher uncertainty calculated by propagating the errors in the detection limit and the counting statistics obtained with the NanoSIMS.

30. Pouchou, J.-L. & Pichoir, F. in *Electron Probe Quantitation* (eds Heinrich, K. F. J. & Newberry, D. E.) 31–75 (Plenum, New York, 1991).
31. Hauri, E. H. *et al.* SIMS analysis of volatiles in silicate glasses: 1. Calibration, matrix effects and comparisons with FTIR. *Chem. Geol.* **183**, 99–114 (2002).
32. Koga, K., Hauri, E. H., Hirschmann, M. & Bell, D. Hydrogen concentration analyses using SIMS and FTIR; comparison and calibration for nominally anhydrous minerals. *Geochim. Geophys. Geosyst.* **4**, doi:10.1029/2002GC000378 (2003).
33. Aubaud, C., Hauri, E. H. & Hirschmann, M. M. Hydrogen partition coefficients between nominally anhydrous minerals and basaltic melts. *Geophys. Res. Lett.* **31**, doi:10.1029/2004GL021341 (2004).

LETTERS

Quadrupole transitions revealed by Borrmann spectroscopy

Robert F. Pettifer^{1,†}, Stephen P. Collins^{2,1} & David Laundry³

The Borrmann effect^{1,2}—a dramatic increase in transparency to X-ray beams—is observed when X-rays satisfying Bragg's law diffract through a perfect crystal. The minimization of absorption seen in the Borrmann effect has been explained by noting that the electric field of the X-ray beam approaches zero amplitude at the crystal planes, thus avoiding the atoms. Here we show experimentally that under conditions of absorption suppression, the weaker electric quadrupole absorption transitions are effectively enhanced to such a degree that they can dominate the absorption spectrum. This effect can be exploited as an atomic spectroscopy technique; we show that quadrupole transitions give rise to additional structure at the L_1 , L_2 and L_3 absorption edges of gadolinium in gadolinium gallium garnet, which mark the onset of excitations from $2s$, $2p_{1/2}$ and $2p_{3/2}$ atomic core levels, respectively. Although the Borrmann effect served to underpin the development of the theory of X-ray diffraction^{3–6}, this is potentially the most important experimental application of the phenomenon since its first observation seven decades ago. Identifying quadrupole features in X-ray absorption spectroscopy is central to the interpretation of 'pre-edge' spectra, which are often taken to be indicators of local symmetry, valence and atomic environment⁷. Quadrupolar absorption isolates states of different symmetries to that of the dominant dipole spectrum, and typically reveals orbitals that dominate the electronic ground-state properties of lanthanides and $3d$ transition metals, including magnetism. Results from our Borrmann spectroscopy technique feed into contemporary discussions regarding resonant X-ray diffraction⁸ and the nature of pre-edge lines identified by inelastic X-ray scattering⁷. Furthermore, because the Borrmann effect has been observed in photonic materials, it seems likely that the quadrupole enhancement reported here will play an important role in modern optics.

The observation of the Borrmann effect^{1,2}, first reported in 1941, provided an unambiguous justification for the so-called dynamical theory of X-ray diffraction^{3–6}, which is applicable to perfect or near-perfect crystals. This theory proceeded from two separate but equivalent theoretical viewpoints, one based on the explicit solution of Maxwell's equations^{9,10} and the other on the multiple scattering from planes of atoms¹¹.

In the present experiment, carried out in the 'symmetric Laue' geometry, we adopted a plate-like crystal which was thick enough to absorb almost the entire incident X-ray beam. Close to the Bragg diffraction condition, we observed the Borrmann effect, or 'anomalous transmission', in which two beams of identical intensities emerge from the opposite face of the crystal: one corresponds to the undeflected beam and the other corresponds to the Bragg reflected beam. Within the crystal there are two solutions to Maxwell's equations, differing only in the phase relationship between the two plane waves. One solution, called the β branch, produces an electric field

that reaches a maximum at the atomic positions, whereas for the other solution, the α branch, the field at the atoms is almost zero. As the dipole absorption is directly proportional to the electric field at the atomic sites, this results in an absorption minimum for the α branch and an absorption maximum for the β branch. The α branch survives even for thick crystals where the normal X-ray transmission is much less than one (typically less than 0.0001 (ref. 10)), and the material exhibits a dramatic increase in transparency.

The electric field can be written (see Fig. 1 and Supplementary Information) as

$$E(\mathbf{r}, t) = 2\omega A \hat{y} e^{i(k_z z - \omega t)} \sin(k_x x)$$

where A is the vector potential amplitude; ω is the angular frequency of the X-ray; $k_x = \pi/d$, where d is the spacing between planes; and $k_z = (2\pi/\lambda)\cos\theta$, where λ is the wavelength of the X-ray inside the material and θ is the Bragg angle. From this equation we can see that the electric field goes to zero at the atomic planes (positioned at $x = nd$, where n is an integer), producing a standing wave in the x direction and a running wave in the z direction. This distribution of electric field is identical to that found between two perfectly conducting sheets operating in the TE_0 waveguide mode¹². In reality, the electric field at the atoms does not vanish completely, owing to crystal imperfections and thermal motion, but the central result—that the electric dipole absorption is greatly reduced—remains valid.

The key idea behind the present work is that although the Borrmann effect minimizes the electric field at atomic sites, it simultaneously maximizes the electric field gradient, which drives electric quadrupole transitions. The net result is a very large relative enhancement of the, usually weak, quadrupole absorption, to such an extent that it can dominate the absorption spectrum. The magnetic field is also non-vanishing at the atomic sites, but in the X-ray regime the contribution of magnetic dipole absorption is negligible¹³ (fortunately, magnetic properties are conveniently accessible through electric transitions¹⁴). We can thus exploit the unusual wave field in the X-ray Borrmann effect to produce a very large enhancement in quadrupole absorption.

Selection rules in X-ray spectroscopy dictate that only electronic states of specific symmetries (angular momenta) are observed. For the dominant electric dipole (E1) transitions, the change in angular momentum is $\Delta l = \pm 1$. For example, L_2 and L_3 absorption, which excite a $2p$ ($l = 1$) core electron are only sensitive to s and d states, whereas the $4f$ electrons that are responsible for most of the interesting properties of lanthanides, including magnetism, can be seen from electric quadrupole (E2) transitions, for which $\Delta l = 0, \pm 2$.

Traditionally, quadrupole transitions are identified using their angular dependence, but in many cases we expect the angular dependence to be weak¹³ (see Supplementary Information). We note

¹Department of Physics, University of Warwick, Coventry CV4 7AL, UK. ²Diamond Light Source Ltd, Diamond House, Harwell Science and Innovation Campus, Didcot OX11 0DE, UK.

³STFC Daresbury Laboratory, Daresbury, Warrington WA4 4AD, UK.

[†]Deceased.

also that E2 effects in photoemission have been studied recently¹⁵ by means of the X-ray standing-wave technique, which relies on the same wave field cancellation in the two-beam case as reported here. There are also interesting parallels between the present work and diffraction anomalous fine structure, which uses resonant scattering to obtain site-specific spectroscopic information. The resonant Borrmann effect can be seen as being quite distinct from other approaches, in that the effect permits only a single, well-defined, wave field solution, which automatically leads to a strong quadrupole enhancement. Moreover, the scattered signals are simple to interpret, as they follow the X-ray absorption, which is related to the imaginary part of the forward-scattering amplitude by means of the optical theorem. Pioneering experiments on the use of two-beam techniques in Compton scattering¹⁶ have been shown to yield non-diagonal terms in the momentum density matrix, and two-beam electron-energy-loss experiments have demonstrated sensitivity to the symmetry of core excitons¹⁷. A recent and very exciting development in X-ray spectroscopy is the application of non-resonant inelastic X-ray scattering (NIXS) to the identification of non-dipolar processes. The NIXS energy-loss spectrum resembles conventional absorption spectra but lacks restrictive angular momentum selection rules, and systematic studies with increasing momentum transfer offer the most obvious comparison with the present work¹⁸. The technique, however, is intrinsically inefficient and requires very significant access to the most advanced synchrotron sources and instrumentation.

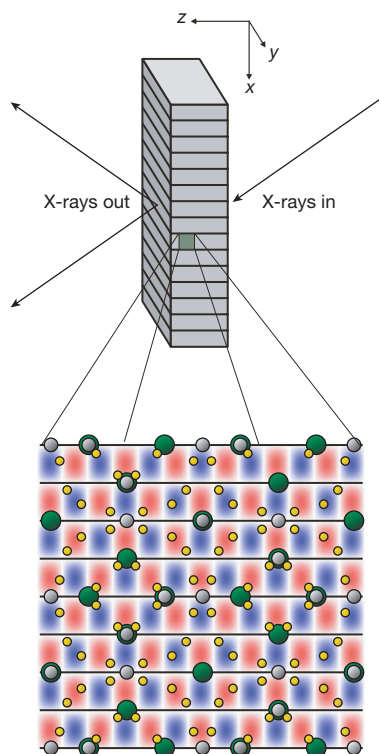


Figure 1 | The Borrmann effect in a gadolinium gallium garnet crystal. An X-ray beam enters the plate-like crystal from the right. Within the crystal a standing wave is formed perpendicular to, and a running wave formed parallel to, the atomic planes by a superposition of the incident and reflected waves. The magnified view shows the electric field (positive amplitude, red; negative amplitude, blue), which propagates from right to left, avoiding all the strongly absorbing gadolinium (green) and gallium (grey) atoms and thus exhibiting ‘anomalous transmission’: two beams of equal intensity emerge from the left of the crystal. Enhanced quadrupole absorption arises because the gradient of the field, which drives this process, is maximized at the atomic planes, where the field vanishes.

It has been shown that the Borrmann effect can be used to detect quadrupole transitions¹⁹, but this was discovered at a time when tuneability of synchrotron radiation was not available; as we report here, it is now possible to study the Borrmann effect as the energy is varied through an absorption edge.

Conventional absorption measurements at the gadolinium L₃ edge in gadolinium gallium garnet, combined with a deconvolution technique²⁰ to partially remove lifetime broadening, reported a weak but well-resolved peak 7.5 eV below the main absorption maximum. The integrated area of this peak (assumed to be quadrupolar) was found to be 2.6% of that of the main edge peak created by E1 absorption (that is, to the empty 5*d* states), which is significant but too small to be resolved by direct measurement.

A detailed analysis (see Supplementary Information) shows that by recording I/I_0 , where I_0 is an intensity measurement of the incident flux and I is the Borrmann intensity, it is possible to measure an effective absorption spectrum with a strongly enhanced quadrupole contribution. Thus, for a perfect crystal

$$\exp(-\mu_A t') = \frac{I}{I_0} = \frac{1}{4} \exp[k_1 - t'(k_2 + \mu_{E2} + \mu_{E1}(1 - \varepsilon))]$$

where μ_A is an apparent linear absorption coefficient, k_1 is a wave-length-dependent parameter that controls the fraction of incident X-rays that participate in the Borrmann effect, k_2 is a parameter that controls absorption by (weakly absorbing) atoms at positions other than the equidistant nodal planes, μ_{E1} and μ_{E2} are respectively the electric dipole (E1) and quadrupole (E2) absorption coefficients under study, ε is a temperature-dependent parameter that is very close to unity¹⁰ (becoming exactly unity in the limiting case in which the atoms lie precisely at the nodes of the wave field), and t' is the effective thickness of the crystal.

Details of the measurements of Borrmann spectroscopy in cubic²¹ gadolinium gallium garnet (Fig. 1), carried out at beamline 16.3 (ref. 22) of the Synchrotron Radiation Source, Daresbury Laboratory, UK, are given in the Methods.

Figure 2a shows the Borrmann absorption at $T = 10$ K, as a function of energy near the gadolinium L₃ edge, along with a conventional absorption spectrum from a thin powdered sample foil. A cursory comparison of the two spectra show the presence of an extra sharp, intense peak in the Borrmann case, located by peak-fitting to be 7.2 eV below the main peak, in good agreement with the position of the pre-edge feature uncovered by deconvolution, and close to the value of 8 eV reported by resonant inelastic X-ray scattering²³. Without further analysis, this pre-edge peak can be unambiguously identified as an electric quadrupole transition to the narrow, half-filled 4*f* states. The Borrmann and conventional data are compared by adjusting the scales to force agreement below the edge and at the peak of the spectrum. This approach assumes that the quadrupole transitions affect mainly the pre-edge region. Because, in the case of strongly enhanced quadrupole absorption, the lineshapes above the edge show some discrepancy with the conventional spectrum, this assumption may be called into question and we might expect to be able to isolate previously unknown quadrupole contributions to the main spectrum.

Figure 2b is the same measurement of the diffracted beam taken at $T = 300$ K. We still see a resolved quadrupole peak, but it is now weaker owing to the increased thermal motion, which places atoms in regions of higher electric field intensity and thus allows dipole absorption to increase. Figure 2c is an L₂-edge measurement also showing considerable, but weaker, quadrupole enhancement, with the quadrupole peak located 6.8 eV below the dipole peak. Given that both the dipole and quadrupole peak areas in gadolinium are expected to scale with the ‘statistical’ value of two between the L₃ and L₂ edges, the fact that their ratios (0.46 and 0.27, respectively) differ significantly raises some interesting questions that will be addressed in future experimental and theoretical studies.

The most surprising result of the present study is the observation of a very large quadrupole peak at the L_1 edge (Fig. 2d), which is the strongest peak in the Borrmann absorption spectrum. Although this peak is coincident with the shoulder, just visible in the conventional spectrum, such a feature had previously been explained in terms of dipolar absorption²⁴ in the related material gadolinium metal. The selection rule for quadrupole transitions, $\Delta l = 0, \pm 2$, means that the Borrmann effect enhances the same d -state absorption that leads to a strong (dipole) peak at the normal L_2 and L_3 edges. Indeed, our measurements suggest that quadrupole transitions should be taken into account for calculations of L_1 edges of all materials with $Z > 58$, as these will have a peak in the d -state density of states at the edge and the gradient of the electric field is large enough to promote E2 transitions.

The quadrupole peak in the L_3 spectrum is unusually sharp, and is found to be well described by a lorentzian lineshape of width ~ 3.6 eV, which is marginally narrower than the calculated²⁵ decay width of ~ 3.8 eV. Given the instrumental broadening of ~ 1 eV, such a narrow lineshape indicates not only a very weak multiplet splitting, as observed in soft X-ray spectroscopy and atomic calculations²⁶, but rules out strong broadening effects, such as those observed recently in

cerium compounds using NIXS¹⁸ and attributed to electronic hybridization with neighbouring oxygen atoms.

It is clear that the Borrmann effect can be used to identify quadrupole features suspected in the pre-edge region. Not only does this effect allow the observation of the isotropic part of the quadrupole absorption, as it does not rely on angle dependence, but it provides an unambiguous identification of quadrupole absorption because only the quadrupole part is enhanced.

Quadrupole transitions play an even more prominent role in types of spectroscopy that are directly sensitive to subtle but important ground-state properties such as orbital ordering²⁷ and magnetism. The problem of extracting magnetic moments from weak dichroism signals observed in hard X-ray spectroscopy, a long-sought goal, has proved difficult, largely because of the poorly understood admixture of dipole and quadrupole contributions. This viewpoint has recently gained credibility from studies of resonant X-ray scattering, where the quadrupole signal is reduced almost to zero by a careful choice of experimental geometry²⁸. Borrmann spectroscopy should significantly aid the development of theories that underpin X-ray studies of magnetism.

A study of the effect of atomic displacement on Borrmann spectroscopy has shown that different temperature factors are required for dipole and quadrupole absorption, and that the quadrupole enhancements illustrated here are not dissimilar to those of a perfect crystal. This will be the subject of another publication. Although the Borrmann effect does rely on at least part of the sample crystal being of very high quality, it will naturally pick out such regions, so the requirement of overall crystal perfection is not so limiting.

We thus see that Borrmann spectroscopy provides a first example of multipole contributions to X-ray absorption being strongly modified by an unusual (non-plane-wave) field. This technique is extremely simple, but it is by no means the only way of carrying out such measurements. We note, for example, that the leading edge of a perfect-crystal Bragg diffraction curve corresponds to the excitation of the α branch and should lead to a similar quadrupole enhancement. Experiments with wave fields that are described by linear combinations of more than two plane waves (including multibeam diffraction) are sure to lead to interesting new results.

One of the most compelling aspects of the present work is that the quadrupole enhancement in the Borrmann effect, unexploited for seven decades, may find application in a very much broader field than that described here. Optical spectroscopy from a highly reflective mirror surface, characterized by wave field nodes close to the sampled volume, should exhibit similar spectral modification. Also, the Borrmann effect and perfect-crystal diffraction is observed in photonic crystals, which are increasingly finding application in communications technology. Photonic crystals should therefore also be sensitive to quadrupole transitions^{29,30}, which may be used in populating long-lived states useful for laser applications.

METHODS SUMMARY

The sample was a plate of cubic gadolinium gallium garnet ($\text{Gd}_3\text{Ga}_5\text{O}_{12}$; space group, $Ia\bar{3}d$), 0.57 mm thick, cut parallel to the (100) planes. The tabulated gadolinium L edges are at 7,243 eV (L_3), 7,930 eV (L_2) and 8,376 eV (L_1). As the photoelectric attenuation, μt , around these edges is of order 100, the measurements were analysed within the 'thick-crystal' regime appropriate for the Borrmann effect. The incident wavelength was selected using the channel-cut Si(111) monochromator that is part of the diffraction beamline 16.3 at the Synchrotron Radiation Source, Daresbury Laboratory. Extracting the X-ray beam from the vertical centre of the synchrotron radiation fan ensured linear polarization perpendicular to the vertical scattering plane. The incident beam intensity was monitored using an ionization chamber in front of the sample, which was mounted in a cryostat with a base temperature of 10 K. Choosing the (008) reflection in symmetric Laue geometry placed nodes at the gadolinium and gallium atoms but not at the oxygen atoms. Fortunately, the absorption coefficient of oxygen at these energies is very small.

The measurements were taken by selecting X-ray energies over a range of values close to the gadolinium L edges and scanning the diffractometer θ and

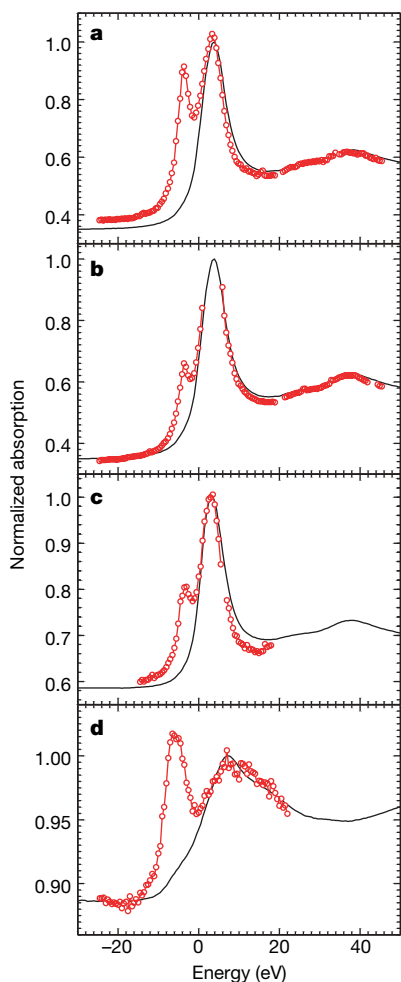


Figure 2 | Comparison of gadolinium absorption spectra in the Borrmann case (red) and the conventional case (black). The scales are adjusted so that the spectra are coincident at the peak in dipole absorption and far from the edge. Energy scales are relative to the values at the absorption edges. **a**, The gadolinium L_3 edge for $T = 10$ K. It can be seen that an additional pre-edge peak exists in the Borrmann data; this represents a quadrupole transition to a $4f$ multiplet. **b**, The gadolinium L_3 edge for $T = 300$ K, showing a weaker quadrupole enhancement. **c**, The gadolinium L_2 edge for $T = 10$ K, again showing an enhanced pre-edge peak. **d**, The gadolinium L_1 edge for $T = 10$ K, exhibiting a strong quadrupole peak that dominates the absorption spectrum.

2θ circles to record the integrated intensity for the Borrmann reflected beam. The transmitted (forward-diffracted) beam was monitored simultaneously using a second ionization chamber to confirm equivalent intensities and lineshapes for the two beams—a characteristic of the (thick-crystal) Borrmann effect¹. The narrow width of reflections remained roughly constant with energy at around 0.005° , indicating that the crystal was of very high quality, although not perfect, as a perfect crystal should exhibit a width that varies with the absorption coefficient¹. An experimental complication was the presence of multiple diffractions in the crystal. These events were eliminated by rotating the crystal by a small amount (up to 5°) about the reciprocal lattice vector.

Received 19 January; accepted 8 May 2008.

- Borrmann, G. The absorbance diagram of quartz. *Phys. Z.* **42**, 157–162 (1941).
- Borrmann, G. Der kleinste Absorptionskoeffizient interferierender Röntgenstrahlung. *Z. Kristallogr. Kristallgeom. Kristallphys. Kristallchem.* **106**, 109–121 (1954).
- Darwin, C. G. The theory of X-ray reflexion. *Phil. Mag.* **27**, 315–333 (1914).
- Darwin, C. G. The theory of X-ray reflexion. Part 2. *Phil. Mag.* **27**, 675–690 (1914).
- Ewald, P. P. Zur Begründung der Kristalloptik. Teil II. *Ann. Phys.* **49**, 117–143 (1916).
- Ewald, P. P. Crystal optics for visible light and X rays. *Rev. Mod. Phys.* **37**, 46–56 (1965).
- De Groot, F. High-resolution X-ray emission and X-ray absorption spectroscopy. *Chem. Rev.* **101**, 1779–1808 (2001).
- Dmitrienko, V. E., Ishida, K., Kirfel, A. & Ovchinnikova, E. N. Polarization anisotropy of X-ray atomic factors and ‘forbidden’ resonant reflections. *Acta Crystallogr.* **A61**, 481–493 (2005).
- Von Laue, M. *Röntgenstrahl-Interferenzen* (Academische, Frankfurt, 1960).
- Batterman, B. W. & Cole, H. Dynamical diffraction of X-rays by perfect crystals. *Rev. Mod. Phys.* **36**, 681–717 (1964).
- Borie, B. The theory of the Borrmann effect in terms of difference equations. *Acta Crystallogr.* **12**, 470–472 (1966).
- Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics* Vol. 2., Ch. 24–10–12 (Addison-Wesley, Reading, Massachusetts, 1964).
- Brouder, C. Angular dependence of X-ray absorption spectra. *J. Phys. Condens. Matter* **2**, 701–738 (1990).
- Lovesey, S. W. & Collins, S. P. *X-ray Scattering and Absorption by Magnetic Materials* (Clarendon, Oxford, UK, 1996).
- Vatanyants, I. A. & Zegenhagen, J. Photoelectric scattering from an X-ray interference field. *Solid State Commun.* **113**, 299–320 (2000).
- Schülke, W. & Mourikis, S. Nondiagonal momentum density of Si by coherent inelastic X-ray scattering. *Acta Crystallogr. A* **42**, 86–98 (1986).
- Batson, P. E. Symmetry-selected electron-energy-loss scattering in diamond. *Phys. Rev. Lett.* **70**, 1822–1825 (1993).
- Gordon, R. A. et al. High multipole transitions in NIXS: Valence and hybridization in 4f systems. *Europhys. Lett.* **81**, 26004 (2008).
- Wagenfeld, H. Normal and anomalous photoelectric absorption of X rays in crystals. *Phys. Rev.* **144**, 216–224 (1966).
- Loeffen, P. W. et al. Deconvolution of lifetime broadening at rare-earth L_{III} edges compared to resonant inelastic X-ray scattering measurements. *Phys. Rev. B* **54**, 14877–14880 (1996).
- Sawada, H. Electron density study of garnets: $Z_3Ga_5O_{12}$; $Z = Nd, Sm, GD, Tb$. *J. Solid State Chem.* **132**, 300–307 (1997).
- Collins, S. P. et al. Station 16.3: a high-resolution single-crystal diffraction facility at the SRS, Daresbury. *J. Synchrotron Rad.* **5**, 1263–1269 (1998).
- Krisch, M. H. et al. Evidence for a quadrupolar excitation channel at the L_{III} edge of gadolinium by resonant inelastic X-ray scattering. *Phys. Rev. Lett.* **74**, 4931–4934 (1995).
- Materlik, G., Müller, J. E. & Wilkins, J. W. L-edge absorption spectra of the rare earths: assessment of the single particle picture. *Phys. Rev. Lett.* **50**, 267–270 (1983).
- Chen, M. H., Crasemann, B. & Kostroun, V. O. Theoretical L_2 - and L_3 -subshell fluorescence yields and L_2 - L_3 X Coster-Kronig transition probabilities. *Phys. Rev. A* **4**, 1–7 (1971).
- Thole, B. T. et al. 3d x-ray-absorption lines and the $3d^9 4f^{n+1}$ multiplets of lanthanides. *Phys. Rev. B* **32**, 5107–5118 (1985).
- Fabrizio, M., Altarelli, M. & Benfatto, M. X-Ray resonant scattering as a direct probe of orbital ordering in transition-metal oxides. *Phys. Rev. Lett.* **80**, 3400–3403 (1998).
- Brown, S. D. et al. Dipolar excitations at the L_{III} X-ray absorption edges of the heavy rare-earth metals. *Phys. Rev. Lett.* **99**, 247401 (2007).
- Zhang, Z. & Satpathy, S. Electromagnetic wave propagation in periodic structures: Bloch wave solution of Maxwell's equations. *Phys. Rev. Lett.* **65**, 2650–2653 (1990).
- Ward, A. J., Pendry, J. B. & Stewart, W. J. Photonic dispersion surfaces. *J. Phys. Condens. Matter* **7**, 2217–2224 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This paper is dedicated to R. F. Pettifer, who died on 5 March 2008, and whose numerous contributions to the field of synchrotron radiation included the original proposal of Borrmann spectroscopy. The authors are grateful to C. Nave and V. E. Dmitrienko for their comments on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.P.C. (steve.collins@diamond.ac.uk).

LETTERS

A multi-component Fermi surface in the vortex state of an underdoped high- T_c superconductor

Suchitra E. Sebastian¹, N. Harrison², E. Palm³, T. P. Murphy³, C. H. Mielke², Ruixing Liang^{4,5}, D. A. Bonn^{4,5}, W. N. Hardy^{4,5} & G. G. Lonzarich¹

To understand the origin of superconductivity, it is crucial to ascertain the nature and origin of the primary carriers available to participate in pairing^{1–6}. Recent quantum oscillation experiments on high-transition-temperature (high- T_c) copper oxide superconductors^{7–10} have revealed the existence of a Fermi surface akin to that in normal metals, comprising fermionic carriers that undergo orbital quantization¹¹. The unexpectedly small size of the observed carrier pocket, however, leaves open a variety of possibilities for the existence or form of any underlying magnetic order, and its relation to d -wave superconductivity^{12–15}. Here we report experiments on quantum oscillations in the magnetization (the de Haas-van Alphen effect) in superconducting $\text{YBa}_2\text{Cu}_3\text{O}_{6.51}$ that reveal more than one carrier pocket. In particular, we find evidence for the existence of a much larger pocket of heavier mass carriers playing a thermodynamically dominant role in this hole-doped superconductor. Importantly, characteristics of the multiple pockets within this more complete Fermi surface impose constraints on the wavevector of any underlying order and the location of the carriers in momentum space. These constraints enable us to construct a possible density-wave model with spiral or related modulated magnetic order, consistent with experimental observations.

$\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ belongs to the family of quasi-two-dimensional copper oxide superconductors that can exhibit surprisingly high T_c s compared to other layered superconductors¹⁶. Until recently⁷, closed Fermi surface sheets in these high- T_c superconductors had posed a considerable challenge to observe, requiring very strong pulsed magnetic fields. In the present experiment we utilize the magnetic torque technique^{17,18} in continuous magnetic fields, which is particularly favourable given the quasi-two-dimensional structure of these materials, enabling us to measure de Haas-van Alphen oscillations under carefully controlled experimental conditions. Figure 1 shows an example of such a measurement on a single crystal of underdoped $\text{YBa}_2\text{Cu}_3\text{O}_{6.51}$ (a sample grown at the University of British Columbia in which quantum oscillations are reported in ref. 7).

The key finding of the present study is the observation of a new oscillatory component of higher frequency $F_\beta = 1,650 \pm 40$ T but almost 30 times smaller in amplitude than the dominant component of frequency first observed in ref. 7 (denoted as F_α in this work; Fig. 2). The comparably weak oscillation amplitude of frequency F_β associated with a second pocket (β) belies its greater thermodynamic significance—an important consideration in the determination of the carriers most relevant in pairing in this quasi-two-dimensional material. The effective quasiparticle mass of carriers in the β pocket, $m_\beta^* = 3.8 \pm 0.4 m_e$ (where m_e is the electron mass), is approximately twice the effective quasiparticle mass (m_α^*) of carriers in the α

pocket—obtained from a temperature-dependent fit to the Lifshitz-Kosevich theory¹¹ (Fig. 3). This new pocket also contains 3.29 ± 0.15 times as many carriers as the α pocket, as determined from the Onsager expression¹¹, $A_k = 2\pi e F / \hbar$, relating the cross-sectional area in momentum space (A_k) to the de Haas-van Alphen frequency F (here e is the electron charge).

Were the α and β pockets to constitute the entire Fermi surface, their total carrier density would be expected to correspond to the nominal hole doping value $p_{\text{nom}} \approx 0.1$ (estimated in ref. 19). The

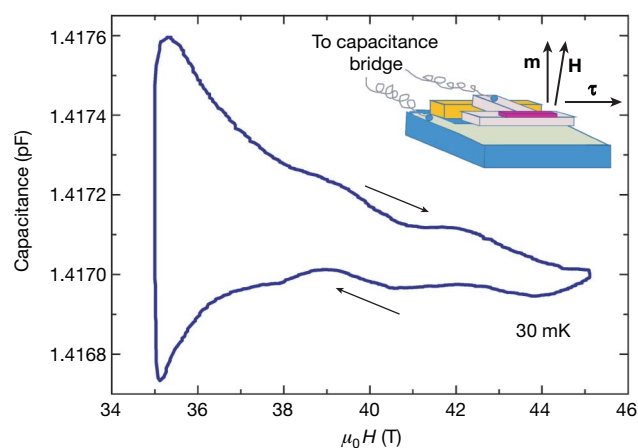


Figure 1 | Experimental data. Capacitance versus magnetic field obtained on sweeping the magnetic field up and down between 35 T and 45 T at base temperature of the dilution refrigerator ($T \approx 30$ mK). The sample is placed on a cantilever whose deflection resulting from the magnetic torque $\tau = \mathbf{m} \times \mathbf{H}$ (due to a small inclination of the magnetic moment \mathbf{m} with respect to the magnetic field \mathbf{H}) is measured capacitively. The angle of inclination θ is minimized to within 3° of the crystalline c axis, such that the angular deflection $d\theta$ of the cantilever due to quantum oscillations is of the order of $0^\circ 0' 1''$. At zero magnetic field, the capacitance is 1.4358 pF. The field-induced change consists of the superposition of quantum oscillatory components, dominated by the known component (α) of frequency $F_\alpha = 500 \pm 20$ T (this work: slightly lower than refs 7 and 10; see Supplementary Information) and quasiparticle effective mass $m_\alpha^* = 1.9 \pm 0.1 m_e$ (ref. 7), and irreversible superconducting components, establishing that quantum oscillations occur in a regime where some level of vortex pinning prevails (details in Supplementary Information). Our observation of de Haas-van Alphen oscillations within the mixed state with similar cyclotron orbit sizes and quasiparticle effective masses m^* as those in the normal state^{7,10} is reminiscent of other strongly type II superconductors³⁰.

¹Cavendish Laboratory, Cambridge University, JJ Thomson Avenue, Cambridge CB3 0HE, UK. ²National High Magnetic Field Laboratory, Los Alamos National Laboratory, MS E536, Los Alamos, New Mexico 87545, USA. ³National High Magnetic Field Laboratory, Florida State University, Tallahassee, Florida 32306, USA. ⁴Department of Physics and Astronomy, University of British Columbia, Vancouver V6T 1Z4, Canada. ⁵Canadian Institute for Advanced Research, Toronto M5G 1Z8, Canada.

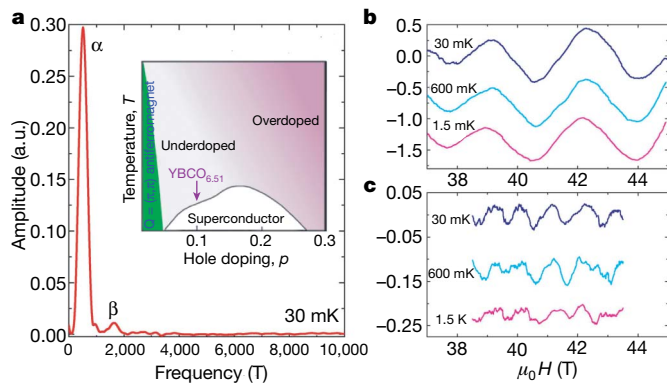


Figure 2 | de Haas-van Alphen oscillations in $\text{YBa}_2\text{Cu}_3\text{O}_{6.51}$. **a**, Fourier transform of six averaged up and down magnetic field sweeps, having performed third order polynomial background subtractions. Here we assume $B = \mu_0 H$. Fourier peaks corresponding to α and β frequencies are seen, as well as a weak feature in the vicinity of 1,000 T that is seen to be consistent with the presence of a small $2F_\alpha$ component when the dominant F_α component is subtracted (see Supplementary Information). Inset, a schematic temperature–doping phase diagram of YBCO. The hole doping is nominally $p_{\text{nom}} \approx 0.1$ for $\text{YBa}_2\text{Cu}_3\text{O}_{6.51}$ (ref. 19). **b**, de Haas-van Alphen oscillations at different temperatures, each averaged from six up and down sweeps after background subtraction. **c**, de Haas-van Alphen oscillations of the new frequency $F_\beta = 1,650 \pm 40$ T after fitting to and subtracting the contribution from the $F_\alpha = 500 \pm 20$ T frequency oscillations and a small 2α plus residual polynomial background (see Supplementary Information), similarly averaged as in **b**. The latter fit and subtraction is performed over a shorter interval in field than **b** (except for the lowest temperature; see Fig. 3b) to eliminate contamination from flux reversal effects upon reversing the magnetic field sweep direction. a.u., arbitrary units.

effective carrier density on assuming Kramer's degeneracy is estimated by the summation $p_{\text{eff}} = 2(n_\alpha A_{k\alpha} + n_\beta A_{k\beta})/A_{\text{BZ}}$ over pocket momentum space areas, where A_{BZ} is the area of the unreconstructed Brillouin zone and n refers to the number of instances each occurs in the reconstructed Brillouin zone (negative for electrons and positive for holes). We find that a model where $n_\alpha = -1$ and $n_\beta = 1$ (such that

$p_{\text{eff}} = 0.082 \pm 0.003$) yields reasonable consistency with p_{nom} , suggesting that the α -pocket is an electron pocket, and the β -pocket a hole pocket.

Given the quasi-two-dimensional electronic structure, the total pocket contribution to the linear temperature coefficient of the specific heat is given by the summation $\gamma_m = \gamma'(|n_\alpha| m_\alpha^* + |n_\beta| m_\beta^*)/m_e$, where $\gamma' = 1.46 \text{ mJ mol}^{-1} \text{ K}^{-2}$ (ref. 8), provided once again that Kramer's degeneracy holds. A single pocket of each type (that is, $|n_\alpha| = 1$ and $|n_\beta| = 1$) would yield $\gamma_m \approx 8.2 \pm 0.6 \text{ mJ mol}^{-1} \text{ K}^{-2}$, comparable in value to $\gamma_C \approx 10 \text{ mJ mol}^{-1} \text{ K}^{-2}$ measured in heat capacity experiments in the normal state above T_c (ref. 20) (details in Supplementary Information).

If we associate Fermi surface reconstruction with the observed pockets, our experimental findings place constraints on the ordering wavevector \mathbf{Q} . We start by making the simplifying single-layer assumption that \mathbf{Q} averages out the effects of bilayer splitting within the $\text{YBa}_2\text{Cu}_3\text{O}_{6.51}$ unit cell by coupling bonding and anti-bonding states. Reconstruction of the Fermi surface shown in Fig. 4a by the commensurate ordering wavevector $\mathbf{Q} = (\pi, \pi)$ (refs 14, 15, 21)—as seen in the Néel phase for $p_{\text{nom}} \leq 0.05$ (ref. 16)—leads to two identical hole pockets at equivalent $\mathbf{k} = (\pi/2, \pi/2)$ points for each electron pocket at $\mathbf{k} = (\pi, 0)$ in the reconstructed Brillouin zone (shown in Fig. 4b). Under these circumstances, where $n_\alpha = -1$ and $n_\beta = 2$, the estimated value of $p_{\text{eff}} = 2(n_\alpha A_{k\alpha} + n_\beta A_{k\beta})/A_{\text{BZ}} \approx 0.20 \pm 0.01$ would be more than twice the nominal doping value $p_{\text{nom}} \approx 0.10$. The observed orbit sizes therefore appear too large to be compatible with a $\mathbf{Q} = (\pi, \pi)$ scenario (Fig. 4b).

We therefore consider a scenario with either a differently commensurate order (similar to that considered in ref. 13) or truly incommensurate order, and show that this could lead to a hole pocket compatible in size with the observed β pocket. We consider the particular case of a single wavevector $\mathbf{Q} = (\pi[1-2\delta], \pi)$, $(\pi[1+2\delta], \pi)$, $(\pi, \pi[1-2\delta])$ or $(\pi, \pi[1+2\delta])$ in $\text{YBa}_2\text{Cu}_3\text{O}_{6.51}$ where δ is taken to be 0.1 (see Fig. 4c). Such a wavevector corresponds to a helical or spiral density-wave scenario in which δ modulates the orientation of the staggered spins within the planes. This value of δ falls within the range seen in inelastic neutron scattering measurements for $p_{\text{nom}} \approx 0.1$ that yield spin fluctuations at

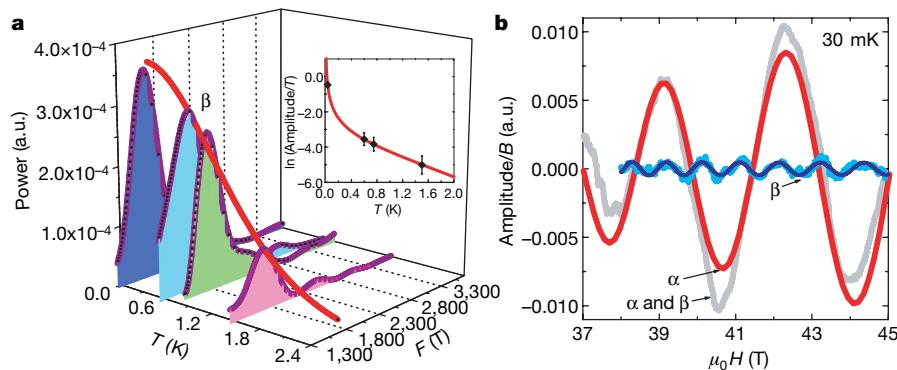


Figure 3 | Fits to the de Haas-van Alphen oscillations. **a**, Power spectrum (amplitude squared, A_T^2) of the averaged β frequency oscillations in Fig. 2c plotted versus temperature T , accompanied by a Lifshitz-Kosevich fit (red line). Inset, a fit of the amplitude divided by temperature (A_T/T) to the Lifshitz-Kosevich temperature dependence: $A_T/T = A_0/\sinh X$, where $X = 14.69 m^* T/B$ (m^* relative to the free electron mass m_e). An effective mass of the β pocket is obtained as $m_\beta^* = 3.8 \pm 0.4 m_e$. The functional form of A_T is derived from the Fermi-Dirac distribution function¹¹. Error bars denote the standard error estimated from the noise floor of the Fourier transform. **b**, Magnetic-field-dependent fits of $A_B/B = A_0 \sin(2\pi F/B + \phi) \exp(-\zeta/B)$ (as would be expected for a two-dimensional Fermi surface)¹¹ to the measured oscillations (in grey and cyan) of each frequency at $T \approx 30$ mK. The red and blue lines show the fit F_α and F_β oscillatory components. The full fit including a polynomial background is shown in the Supplementary Information. For the α pocket, we obtain

$\zeta_\alpha = 160 \pm 50$ T, for which possible interpretations in terms of the scattering rate are discussed in the Supplementary Information. The frequency ratio $F_\beta/F_\alpha \approx 3.29 \pm 0.15$ suggests that the β frequency is not an intrinsic harmonic of the α frequency. As the exponential damping factor scales with the harmonic index, the large value for ζ_α also implies that the $3F_\alpha$ harmonic of the α frequency should be unobservably small in the Lifshitz-Kosevich theory¹¹. Other potential sources of nonlinearities that could lead to harmonics are discussed in the Supplementary Information—ultimately, a distinction between the $3F_\alpha$ harmonic of the α frequency thus generated and the β frequency will depend on the precision to which the frequencies and the relative phase between the oscillations can be measured (see Supplementary Information). Although the β oscillations shown in **b** are consistent with a fit periodic in $1/\mu_0 H$, it appears that the weaker than expected field-dependence of the amplitude cannot be described by a simple exponential damping factor (see Supplementary Information).

wavevectors $\mathbf{Q} = (\pi[1 \pm 2\delta], \pi)$ (ref. 22), although long-range order has not been observed at zero magnetic field in $\text{YBa}_2\text{Cu}_3\text{O}_{6.51}$ (ref. 23). In such a scenario, the hole pockets at equivalent $\mathbf{k} = (\pi/2, \pi/2)$ points would be significantly asymmetric in size. In the limit of a sufficiently large ordering gap or δ , only a single hole pocket would appear for each electron pocket (such that $n_\beta = 1$ instead of 2). In this case, the effective halving of the number of hole pockets with respect to the commensurate case results in a doubling of the hole pocket size, thereby reducing the density of states. Pocket sizes consistent with our experimental findings can arise from such a Fermi surface reconstruction (diagram in Fig. 4d), yielding a value of $p_{\text{eff}} = 0.082$, which is reasonably close to p_{nom} , with a similar agreement for γ_C .

Collinear structures, on the other hand (in which the amplitude of the spin is modulated by δ), would involve multiple $\pm\mathbf{Q}$ vectors (as recently proposed for $\delta = 1/8$ in ref. 13) and in general lead to a hierarchy of gaps and smaller or no hole pockets^{13,24}. In this case, prominent orbits like the β orbit in Fig. 4d could become observable as a result of magnetic breakdown tunnelling overcoming all but the dominant gap in strong magnetic fields, as occurs in elemental Cr (ref. 24). These scenarios, with a different commensurability or incommensurability, may be equally applicable to all forms of density wave order^{2,4,14,15}.

The bilayer potential—the effect of which was considered small in the simplified single-layer scenario considered thus far—may, however, play an important role in determining a complete picture of the

Fermi surface. A significant bilayer potential could lift the assumed degeneracy of either or both pockets, leading to dual frequencies for the affected pocket, one of which could even be completely depopulated (more likely for the electron pocket, given its smaller size). It might then be possible that the observed α and β frequencies each correspond to split frequencies for either or both electron and hole pockets.

We conclude with a brief discussion of further issues raised by this study. The first concerns the origin of the potential Fermi surface reconstruction in relation to superconductivity. One possibility is that modulated magnetic ordering competes with superconductivity, nucleating wherever superconductivity is suppressed, as in the vortex core regions. In the case of antiferromagnetic order, this behaviour may have been identified in heavy fermion compounds²⁵, and has been suggested by neutron scattering on $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ (ref. 26) and muon spin rotation on $\text{YBa}_2\text{Cu}_3\text{O}_{6.5}$ (ref. 27). Suppression of antiferromagnetism by the onset of superconductivity may account for the loss of neutron scattering intensity at the incommensurate wavevectors at low energies and low temperatures observed in $\text{YBa}_2\text{Cu}_3\text{O}_{6.5}$ (ref. 23).

The second issue concerns the comparative stability of helical (or spiral)²⁸ versus collinear forms of order^{3,13}, which could affect the coexistence of superconductivity with magnetic order. Whereas the former favours a uniform distribution of doped-holes²⁸, the latter is conducive to formation of stripes³. In either scenario, magnetic-field-induced Landau quantization consistent with the observation of quantum oscillations could be important in stabilizing long-range order or tuning \mathbf{Q} in strong magnetic fields, as observed in other layered metals²⁹.

A third issue concerns the renormalization of the cyclotron mass for the two Fermi surface pockets identified. A parameterization of the unreconstructed band structure for $\text{YBa}_2\text{Cu}_3\text{O}_{6.51}$, such as that adopted in ref. 13, predicts pockets of similar unrenormalized band masses. The higher mass renormalization of hole pockets suggested by our observations would, therefore, need to be taken into consideration within any explanation of superconductivity in this system.

Note added in proof: A potential commensurate ordering scenario that could explain a large hole orbit has recently been proposed³¹.

Received 21 December 2007; accepted 7 May 2008.

1. Damascelli, A., Hussain, Z. & Shen, Z. X. Angle-resolved photoemission studies of the cuprate superconductors. *Rev. Mod. Phys.* **75**, 473–541 (2003).
2. Lee, P. A. From high temperature superconductivity to quantum spin liquid: Progress in strong correlation physics. Preprint at (<http://arXiv.org/abs/0708.2115v2>) (2007).
3. Kivelson, S. A. *et al.* How to detect fluctuating stripes in the high-temperature superconductors. *Rev. Mod. Phys.* **75**, 1201–1241 (2003).
4. Chakravarty, S., Laughlin, R. B., Morr, D. K. & Nayak, C. Hidden order in the cuprates. *Phys. Rev. B* **63**, 094503 (2001).
5. Anderson, P. W. *et al.* The physics behind high-temperature superconducting cuprates: The 'plain vanilla' version of RVB. *J. Phys. Condens. Matter* **16**, R755–R769 (2004).
6. Kaul, R., Kim, Y. B., Sachdev, S. & Senthil, T. Algebraic charge liquids. *Nature Phys.* **4**, 28–31 (2008).
7. Doiron-Leyraud, N. *et al.* Quantum oscillations and the Fermi surface in an underdoped high- T_c superconductor. *Nature* **447**, 565–568 (2007).
8. Yelland, E. A. *et al.* Quantum oscillations in the underdoped cuprate $\text{YBa}_2\text{Cu}_4\text{O}_8$. *Phys. Rev. Lett.* **100**, 047003 (2007).
9. Bangura, A. F. *et al.* Small Fermi surface pockets in underdoped high temperature superconductors: Observation of Shubnikov–de Haas oscillations in $\text{YBa}_2\text{Cu}_4\text{O}_8$. *Phys. Rev. Lett.* **100**, 047004 (2007).
10. Jaudet, C. *et al.* de Haas–van Alphen oscillations in the underdoped cuprate $\text{YBa}_2\text{Cu}_3\text{O}_{6.5}$. Preprint at (<http://arXiv.org/abs/0711.3559>) (2007).
11. Shoenberg, D. *Magnetic Oscillations in Metals* (Cambridge Univ. Press, Cambridge, UK, 1984).
12. LeBoeuf, D. *et al.* Electron pockets in the Fermi surface of hole-doped high- T_c superconductors. *Nature* **450**, 533–536 (2007).
13. Millis, A. J. & Norman, M. Antiphase stripe order as the origin of electron pockets observed in 1/8-hole-doped cuprates. *Phys. Rev. B* **76**, 220503(R) (2007).
14. Chen, W.-O., Yang, K.-Y., Rice, T. M. & Zhang, F. C. Quantum oscillations in magnetic field induced antiferromagnetic phase of underdoped cuprates: Application to ortho-II $\text{YBa}_2\text{Cu}_3\text{O}_{6.5}$. *Europhys. Lett.* **82**, 17004–17008 (2008).

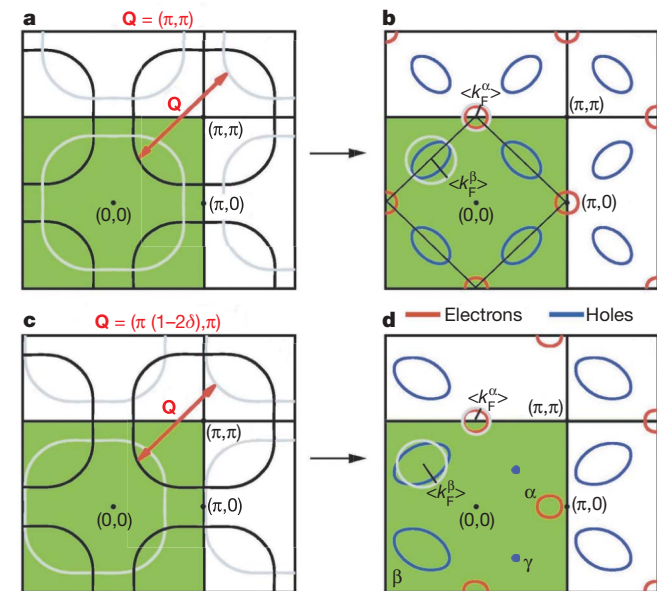


Figure 4 | Fermi surface reconstruction in $\text{YBa}_2\text{Cu}_3\text{O}_{6.51}$. **a**, Schematic Fermi surface reconstruction for a commensurate ordering wavevector $\mathbf{Q} = (\pi, \pi)$ and nominal doping $p_{\text{nom}} = 0.1$ in the extended Brillouin zone representation (unit cell shaded in green). The original unreconstructed Fermi surface (black) represents that for $\text{YBa}_2\text{Cu}_3\text{O}_{6.51}$ (details in Supplementary Information). Translation by commensurate \mathbf{Q} (grey) shifts the original dispersion to $\epsilon_{\mathbf{k}+\mathbf{Q}}$, leading to hole pockets at $(\pi/2, \pi/2)$, and possibly electron pockets at $(\pi, 0)$, depending on the magnitude of the gap opening between $\epsilon_{\mathbf{k}}$ and $\epsilon_{\mathbf{k}+\mathbf{Q}}$. **b**, The hole (blue) and electron (red) pockets expected from **a** for a nominal doping $p_{\text{nom}} = 0.1$. **c**, Representation of a different translation vector $\mathbf{Q} = (\pi[1-2\delta], \pi)$ assuming $\delta \approx 0.1$, corresponding to a simple helical or spiral spin-density wave scenario (see Supplementary Information). **d**, Hypothetical Fermi surface on translation by single \mathbf{Q} shown in **c** ($\pm\mathbf{Q}$ translation shown in Supplementary Information). Good agreement with the observed pockets is obtained on using $p_{\text{eff}} = 0.082$. For comparison, the respective sizes of the pockets measured experimentally from the average Fermi wavevector $\langle k_F \rangle = \sqrt{A_k/\pi}$ are represented graphically by grey circles in **b** and **d**. A reduced δ or ordering gap (or an increased p_{eff}) could potentially lead to an additional small hole pocket labelled as γ (see Supplementary Information).

15. Chakravarty, S. & Kee, H.-Y. Fermi pockets and quantum oscillations of the Hall coefficient in high temperature superconductors. Preprint at (<http://arXiv.org/abs/0710.0608>) (2007).
16. Schrieffer, J. R. & Brooks, J. S. (eds) *High-Temperature Superconductivity Theory and Experiment* (Springer Science, New York, 2007).
17. Naughton, M. J. *et al.* Magnetization study of the field-induced transitions in tetramethyltetraselenafulvalenium perchlorate, (TMTSF)₂ClO₄. *Phys. Rev. Lett.* **26**, 969–972 (1985).
18. Wosnitza, J. *et al.* Two-dimensional Fermi liquid with a fixed chemical potential. *Phys. Rev. B* **61**, 7383–7387 (2000).
19. Liang, R., Bonn, D. A. & Hardy, W. N. Evaluation of CuO₂ plane hole doping in YBa₂Cu₃O_{6+x} single crystals. *Phys. Rev. B* **73**, 180505(R) (2006).
20. Loram, J. W., Luo, J., Cooper, J. R., Liang, W. Y. & Tallon, J. L. Evidence on the pseudogap and condensate from the electronic specific heat. *J. Phys. Chem. Solids* **62**, 59–64 (2001).
21. Kampf, A. Magnetic correlations in high temperature superconductivity. *Phys. Rep.* **249**, 219–351 (1994).
22. Stock, C. *et al.* Dynamic stripes and resonance in the superconducting and normal phases of YBa₂Cu₃O_{6.5}. *Phys. Rev. B* **69**, 014502 (2004).
23. Stock, C. *et al.* From incommensurate to dispersive spin-fluctuations: The high-energy inelastic spectrum in superconducting YBa₂Cu₃O_{6.5}. *Phys. Rev. B* **71**, 024522 (2001).
24. Fawcett, E. Spin-density-wave antiferromagnetism in chromium. *Rev. Mod. Phys.* **60**, 209–283 (1988).
25. Park, T. *et al.* Hidden magnetism and quantum criticality in the heavy fermion superconductor CeRhIn₅. *Nature* **440**, 65–68 (2006).
26. Lake, B. *et al.* Spins in the vortices of a high-temperature superconductor. *Science* **291**, 1759–1762 (2001).
27. Miller, R. I. *et al.* Evidence for static magnetism in the vortex cores of ortho-II YBa₂Cu₃O_{6.50}. *Phys. Rev. Lett.* **88**, 137002 (2002).
28. Shraiman, B. I. & Siggia, E. D. Spiral phase of a doped quantum antiferromagnet. *Phys. Rev. Lett.* **62**, 1564–1567 (1989).
29. Chaikin, P. M. Field induced spin density waves. *J. Phys. I (France)* **6**, 1875–1898 (1996).
30. Corcoran, R. *et al.* de Haas-van Alphen effect in the superconducting state. *Physica B* **206 & 207**, 534–541 (1995).
31. Podolsky, D. & Kee, H.-Y. Quantum oscillations of ortho-II YBa₂Cu₃O_{6.5}. Preprint at (<http://arXiv.org/abs/0806.0005>) (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the National Science Foundation, the Department of Energy (US), Florida State, the UK EPSRC, the Canadian Institute for Advanced Research, and NSERC. S.E.S. acknowledges support from the Institute for Complex Adaptive Matter, COST, and Trinity College (Cambridge University). We acknowledge discussions with E. Abrahams, P. W. Anderson, E. Berg, A. Carrington, S. Chakravarty, J. Fletcher, L. P. Gor'kov, S. R. Julian, H.-Y. Kee, S. A. Kivelson, D. LeBoeuf, P. A. Lee, P. B. Littlewood, A. P. Mackenzie, A. Millis, M. R. Norman, D. Pines, C. Proust, T. M. Rice, S. Sachdev and L. Taillefer, and experimental assistance from G. Jones, J. H. Park and S. Tozer.

Author Contributions Torque experiments were performed by S.E.S., N.H. and G.G.L.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.E.S. (suchitra@phy.cam.ac.uk).

LETTERS

Preseismic velocity changes observed from active source monitoring at the Parkfield SAFOD drill site

Fenglin Niu¹, Paul G. Silver², Thomas M. Daley³, Xin Cheng¹ & Ernest L. Majer³

Measuring stress changes within seismically active fault zones has been a long-sought goal of seismology. One approach is to exploit the stress dependence of seismic wave velocity, and we have investigated this in an active source cross-well experiment at the San Andreas Fault Observatory at Depth (SAFOD) drill site. Here we show that stress changes are indeed measurable using this technique. Over a two-month period, we observed an excellent anticorrelation between changes in the time required for a shear wave to travel through the rock along a fixed pathway (a few microseconds) and variations in barometric pressure. We also observed two large excursions in the travel-time data that are coincident with two earthquakes that are among those predicted to produce the largest coseismic stress changes at SAFOD. The two excursions started approximately 10 and 2 hours before the events, respectively, suggesting that they may be related to pre-rupture stress induced changes in crack properties, as observed in early laboratory studies^{1,2}.

It is well known from laboratory experiments that seismic velocities vary with the level of applied stress^{3–5}. Such dependence is attributed to the opening and closing of microcracks due to changes in the stress normal to the crack surface^{6–8}. In principle, this dependence constitutes a stress meter, provided that the induced velocity changes can be measured precisely and continuously. Indeed, there were several attempts in the 1970s to accomplish this goal using either explosive or non-explosive surface sources^{9–11}. The source repeatability and

the precision in travel-time measurement appeared to be the main challenges in making conclusive observations.

With the availability of highly repeatable sources, modern data acquisition systems and advanced computational capability, Yamamura *et al.*¹² showed compelling evidence that seismic velocity along a baseline in a vault near the coast of Miura Bay, Japan, responds regularly to tidal stress changes. Silver *et al.*¹³ found an unambiguous dependence of seismic velocity on barometric pressure from a series of cross-well experiments at two test sites in California. The stress sensitivity depends primarily on crack density and has a strong nonlinear dependence on confining pressure. Consequently, crack density is expected to decrease rapidly with depth, as should stress sensitivity. It is thus unclear whether the stress-induced velocity variations observed at shallow depths^{12,13} are still detectable at seismogenic depth.

To explore stress sensitivity at seismogenic depth, we have conducted an experiment at Parkfield, California, where adjacent deep wells, namely the SAFOD pilot and main holes (Fig. 1), are available. Accurately located seismicity, together with the availability of high-quality geophysical data in the Parkfield region, make it one of the best areas to detect temporal changes related to the earthquake cycle. A specially designed 18-element piezoelectric source and a three-component accelerometer were deployed inside the pilot and main holes, respectively, at ~1 km depth (see Methods). The experiment was conducted for ~2 months: the first period was 29 October to 28

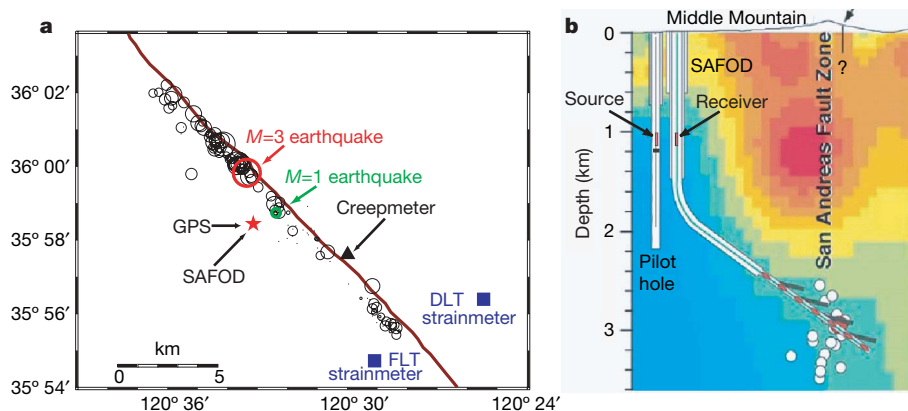


Figure 1 | Map of the experiment site. **a**, Circles show earthquakes that occurred during the experiment period. The $M = 3$ and $M = 1$ events are shown as red and green circles, respectively. Star indicates the Parkfield SAFOD drill site, where the experiment was conducted. Triangle, location of the Middle Mountain creepmeter; squares, locations of the Donalee (DLT) and Frolich (FLT) Gladwin borehole tensor strainmeters. **b**, A vertical section (schematic) of the SAFOD main and pilot holes. Red vertical lines

indicate the source and receiver locations. Background image is electrical resistivity¹⁸ with blue (red) corresponding to relatively high (low) resistivity. White circles show the seismicity, red dashes indicate the SAFOD instrumentation, and black short lines represent sub-horizontal holes drilled off the main hole. The profile at the top of the panel is the surface topography with the arrow indicating the surface exposure of the fault.

¹Department of Earth Science, MS-126, Rice University, 6100 Main Street, Houston, Texas 77005, USA. ²Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, NW, Washington DC 20015, USA. ³Earth Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA.

November 2005, and the second was 11 December 2005 to 10 January 2006. We fired a pulse with a width of 1 ms four times per second and recorded 200-ms-long data with a sampling rate of 48,000 Hz. The waveforms were automatically stacked in groups of 100 shots, resulting in one record (Fig. 2) acquired every 27 s (two additional seconds were needed in storing the data).

To enhance the signal-to-noise ratio (SNR) of the data, we further stacked the raw seismograms in sets of 100. This stacking procedure reduced the data to one stack every 45 min. The 45-min stacked records were then processed with a bandpass filter of 1–5 kHz before the travel-time analysis. We used a cross-correlation-based method to estimate the delay time, which permits subsample precision (see Methods). No smoothing or filtering was applied to the measured delay time series. The error in delay time measurement was estimated to be $\sim 1.1 \times 10^{-7}$ s, based on SNR analysis (see Methods), and this estimate was confirmed by comparing measurements from consecutive recordings. As the nominal travel time of the shear-wave (S-wave) coda along the baseline is about 10 ms, the detectable threshold of velocity perturbation is $\sim 1.1 \times 10^{-5}$, or 11 p.p.m.

We measured the delay times of the S wave and the S wave plus its coda up to 20 ms with respect to a fixed reference trace for each period (Fig. 3). The measurements show daily cycles that are well correlated with the temperature record (Fig. 3). Silver *et al.*¹³ found that this temperature sensitivity originates from the electronics of the recording system rather than from changes in the subsurface velocity field. We excluded the measurements of the first few days to allow the source and sensor to be stabilized at their locations. We also removed the linear trend from the data as was done by Silver *et al.*¹³. In general, the delay times of the coda are about twice as large as those of the S wave, suggesting that they are caused by a change in the velocity of the bulk media, as the coda travels for a longer time in the media and thus is expected to accumulate a larger travel-time anomaly. The delay time closely follows the barometric pressure changes for the first period (Fig. 3a).

After removing the temperature effect from the measured delay time variations (Fig. 3a), we obtained a delay time change of ~ 3.0 μ s in the first period. The corresponding velocity perturbation is about

3×10^{-4} , about an order of magnitude higher than the detectable threshold. During the same time period, the change in barometric pressure is ~ 1.3 kPa. We used a linear regression to estimate the stress sensitivity of the velocity and obtained a value of 2.4×10^{-7} Pa⁻¹. We also calculated the predicted solid Earth tides at the site in the same period and found that the tidal stress varies within 240 Pa, nearly an order of magnitude smaller than changes in barometric pressure. Thus the travel-time changes induced by tidal stress are of the order of 10^{-7} s, close to the measurement error and thus are predicted to be undetectable.

The negative correlation between travel time and barometric stress can be further seen in the delay time data through to the end of the ninth day of the second period. After this time the relationship starts to break down, and we observe instead two prominent excursions in the delay time data that are not seen in the barometric pressure record. It is also confirmed that the two excursions were not caused by precipitation or instrumentation. The amplitudes of the two excursions are ~ 5.5 μ s and ~ 1.5 μ s, respectively, over the nominal ~ 10 ms coda travel time. Using our measured stress sensitivity of 2.4×10^{-7} Pa⁻¹, the corresponding stress changes are 2.3 kPa and 625 Pa for the first and second peak, respectively.

In order to evaluate the possibility of a tectonic cause for the excursions, we examined the seismicity around the SAFOD site occurring in the experiment period (Fig. 4a). The first peak appears to correspond to the largest earthquake occurring in this period (date, 24 December 2005; time, 10:10:57.21 (h:min:s UT); location, 35.9970° N, 120.5565° W; depth, 3.88 km; magnitude 3.00, hereafter the $M = 3$ event), while the second peak corresponds to the second closest (1.5 km) event to the experiment site (29 December 2005, 01:32:50.87 UT, 35.9788° N, 120.5397° W, depth 1.82 km, magnitude 0.98, hereafter the $M = 1$ event). The closest event is about 1.3 km away from our site, but its size is only $M = 0.34$ and thus should not have a large effect at the site.

We calculated the predicted static stress change at SAFOD associated with these two earthquakes. The near-field static displacement at a location r with respect to the earthquake is proportional to $M_0 r^{-2}$, where M_0 is the seismic moment¹⁴. The spatial derivative of

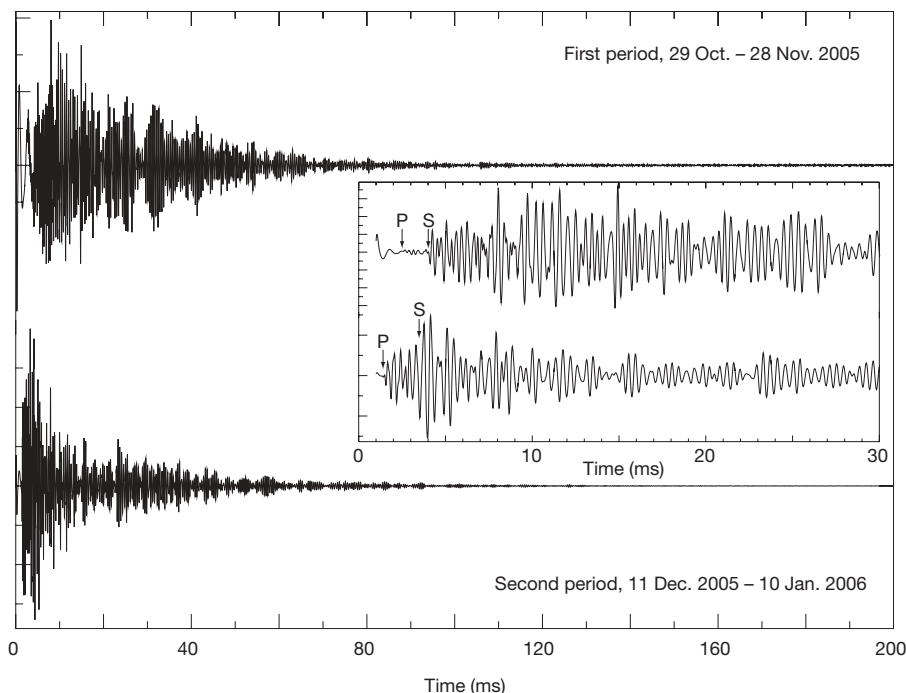


Figure 2 | An example of the raw seismograms obtained from a horizontal component in the two periods. Both are filtered with a band pass filter of

1–5 kHz. Inset shows the first 30 ms of the waveforms. P and S indicate the compressional and shear wave arrivals.

displacement, strain, thus should be $\propto M_0 r^{-3}$. The static stress change at \mathbf{r} is $\Delta\sigma = a \frac{\mu L^2 D}{r^3} = a \frac{\mu(D/L)}{(r/L)^3} = a \frac{\Delta\sigma_0}{\hat{r}^3}$, where $\Delta\sigma_0$ is the average static stress change along the fault, \hat{r} is the characteristic distance measured in fault lengths (L), D is slip on the fault, μ is the shear modulus, and a is a scaling constant equal to $1/(6\pi)$ (ref. 14). If we assume a static stress change in the range of 3 to 10 MPa (refs 15, 16), then the static coseismic stress change at the SAFOD site is estimated to be ~ 250 – 833 Pa for the $M = 3$ event, which is a few times lower than the total stress change (2.3 kPa) calculated from the amount of delay time during the first excursion. The predicted static stress changes at the SAFOD site calculated from the entire local seismicity catalogue are shown in Fig. 4b. Here we used all the events that occurred within 10 km of the site and made a time series of the coseismic stress changes. The $M = 3$ earthquake obviously has the

largest effect at the experiment site. The second largest peak around day 20 corresponds to a relatively deep event (22 November 2005, 03:38:02.13 UT, 36.0100°–120.5692°, depth 5.07 km, $M = 2.6$), which is not observed in the delay time data. The third peak corresponds to the $M = 1$ event. It is not clear to us why the larger $M = 2.6$ event is not observed while the smaller $M = 1$ event shows clearly in the delay time data. But we noticed that data collected in the second period had a better SNR than those of the first period. The associated stress change of the $M = 2.6$ event thus might be below the resolution of the first-month data.

Coseismic change was also observed in other geodetic data. We found a step-function change from the borehole fibre-optic strainmeter data at SAFOD (Fig. 4a inset) as well as from the surface creepmeter data at Middle Mountain (Fig. 4b). The static strain change observed at SAFOD is $\sim (20$ – $25) \times 10^{-9}$, corresponding to

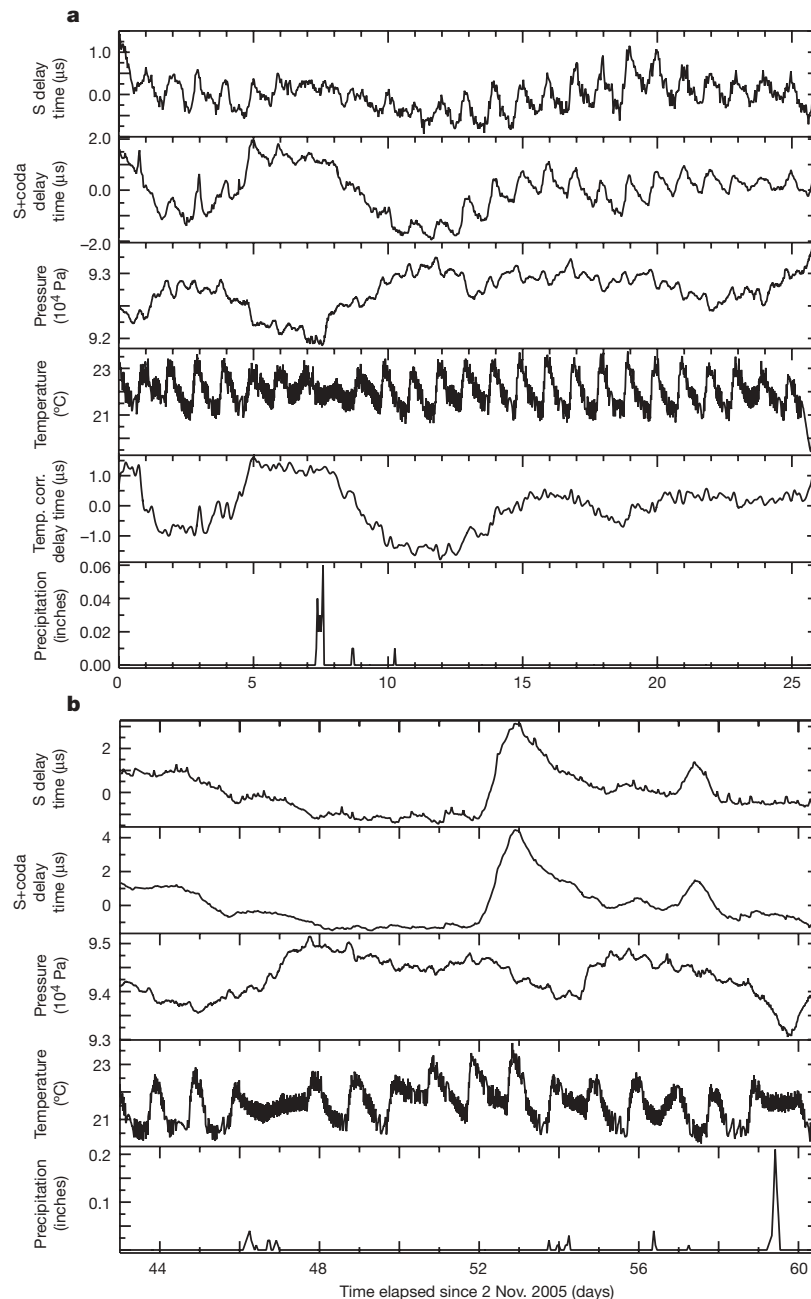


Figure 3 | Estimated delay times for the two periods. **a**, First period; **b**, second period. In each panel, the top two traces are delay times estimated from time windows that contain respectively the S-wave arrival and the S-wave arrival plus the coda; the third trace is barometric pressure; the

fourth trace is temperature; and the bottom trace is precipitation. In **a**, the second to last trace is a temperature corrected version of the second trace. Elapsed time is calculated from 2 November 2005, 00:00:00 UT.

a coseismic stress change of $\sim 600\text{--}750\text{ Pa}$, which is of the same order of magnitude as our estimate. On the other hand, there were no obvious changes in the SAFOD GPS, or the FLT and DLT strainmeter records (Fig 4b). The lack of an observable coseismic signal at these sites is, however, predicted by the theoretical amplitude.

The coseismic offset recorded by the SAFOD strainmeter is not obviously present in the delay time data measured either from the

manually stacked 45-min-per-sample data or from the delay times calculated from the 27-s-per-sample raw data. The derivative of the delay time series (dotted line in Fig. 4c), however, does reveal that the largest offset of the entire two-month observing period occurred $\sim 30\text{ s}$ after the $M = 3$ earthquake. This suggests that there was a small coseismic change in the delay time data. The lack of a stronger coseismic signal in the delay time data may imply that the velocity changes we observed here are mainly the result of a poroelastic¹⁷ rather than an elastic response to abrupt stress changes.

The two travel-time excursions appear to possess significant pre-seismic components. The first excursion was observed to start at 23:34 UT on 23 December 2005, while the $M = 3$ earthquake occurred at $\sim 10.6\text{ h}$ later, at 10:10 UT on 24 December 2005 (Fig. 4c). The excursion reached a maximum right after the earthquake, peaking at 21:21 UT on 24 December 2005. The excursion thus has a clear pre-seismic component besides the coseismic/postseismic changes. The pre-seismic and coseismic/postseismic components account for $\sim 46\%$ and $\sim 54\%$ of the total change, respectively. This is also true for the second excursion. Its onset is around 22:59 UT on 28 December 2005, about 2.5 h before the occurrence of the $M = 1$ earthquake (01:32 UT on 29 December, Fig. 4c).

With the available geodetic instrumentation, it was impossible to further evaluate the pre-seismic component. The most direct test would have been with the SAFOD borehole strainmeter data. Unfortunately, the low frequency component is severely contaminated by surface temperature variations and is unusable for periods longer than a few minutes, and is thus not useful in confirming the two low-frequency excursions (M. Zumberge, personal communication). All other instrumentation is either too far away or not sufficiently sensitive to observe even the coseismic offset. Historically, there has been an absence of pre-seismic signals in geodetic observations, such as a borehole strainmeter. We suggest that this may be the result of two differences between such instruments and our 'stress meter'. First, our basic measurement is not strain, but rather a stress-induced change in the effective elastic constants of a poroelastic medium, mediated by variations in crack properties and fluid flow. These changes may register only weakly on a strainmeter, a GPS, or a creepmeter. Second, a conventional strainmeter measures local change in the volume immediately surrounding the instrument while our measurements reflect stress/strain changes occurring over a volume sampled by the coda waves that could be orders of magnitude larger.

We put forward the hypothesis that there is a change in effective elastic moduli before rupture, such as a sudden increase in microcrack density, which is a phenomenon related to dilatancy and observed in many laboratory studies^{1,2}. As such, further continuous seismic monitoring might provide an effective tool for understanding the stress changes that accompany and perhaps precede seismic activity.

METHODS SUMMARY

We used a specially built piezoelectric source and a Geode recorder (Geometrics) to generate and record seismic waves travelling along a $\sim 10\text{ m}$ baseline near the San Andreas fault at $\sim 1\text{ km}$ depth. A cosine fitting method was used to estimate the S-wave travel time to subsample precision.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 22 October 2007; accepted 12 May 2008.

1. Brace, W. F., Paulding, B. W. & Scholz, C. H. Dilatancy in the fracture of crystalline rocks. *J. Geophys. Res.* **71**, 3939–3953 (1966).
2. Scholz, C. H. Microfracturing and the inelastic deformation of rock I: Compression. *J. Geophys. Res.* **73**, 1417–1432 (1968).
3. Birch, F. The velocity of compressional waves in rocks to 10 kilobars, part 1. *J. Geophys. Res.* **65**, 1083–1102 (1960).
4. Birch, F. The velocity of compressional waves in rocks to 10 kilobars, part 2. *J. Geophys. Res.* **66**, 2199–2224 (1961).
5. Nur, A. & Simmons, G. The effect of saturation on velocity in low porosity rocks. *Earth Planet. Sci. Lett.* **7**, 183–193 (1969).
6. Walsh, J. B. The effect of cracks on the compressibility of rock. *J. Geophys. Res.* **70**, 381–389 (1965).

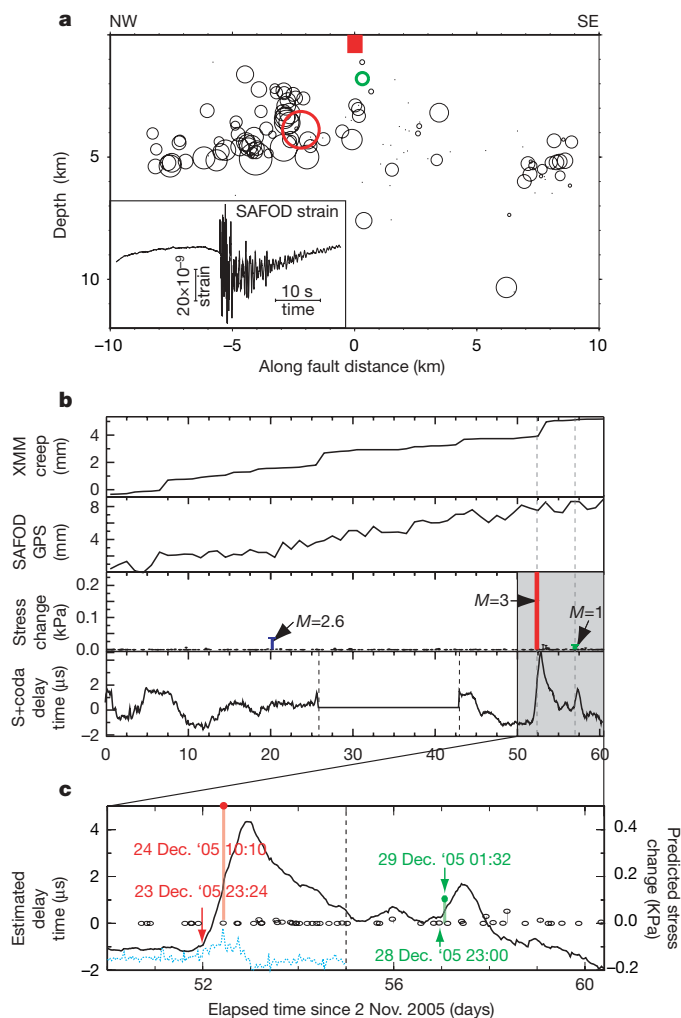


Figure 4 | A comparison of delay time variations with local seismicity and other deformation measurements. **a**, Depth distribution of earthquakes that occurred in the experimental period. Red square, the SAFOD experiment site; red and green circles, the $M = 3$ and $M = 1$ earthquake, respectively. Inset, the SAFOD strainmeter record, which shows a step-function coseismic strain change. The low frequency content of the strainmeter data is severely contaminated by surface temperature variations, and is consequently not suitable for analysis. **b**, Top to bottom: creep measurement at Middle Mountain (XMM); GPS measurement of fault-parallel motion at the SAFOD site; calculated static coseismic stress changes at the SAFOD experiment site for all of the earthquakes; and delay times estimated from the S wave plus its coda for comparison. Dashed lines indicate the time when the $M = 3$ and $M = 1$ earthquakes occurred. Note that the amplitude of the stress change of the $M = 3$ event ($\sim 0.5\text{ kPa}$) is saturated in this plot. **c**, Predicted coseismic stress changes at SAFOD for earthquakes occurring between 22 December 2005 (day 50) and 1 January 2006 (day 60) indicated by shading in **b** are shown with the delay time estimation. Stress changes from the local seismicity between days 55 and 60 are amplified by a factor of 10. The two filled circles show the stress change of the $M = 3$ (red) and $M = 1$ (green) event, respectively. The vertical lines indicate the occurrence times of the $M = 3$ and the $M = 1$ event, and the red and green (upward) arrows show the onset times of the two excursions. Blue dotted line is the derivative of the delay time series. Notice that the largest change occurred about $\sim 30\text{ s}$ after the $M = 3$ earthquake.

7. Nur, A. Effects of stress on velocity anisotropy in rocks with cracks. *J. Geophys. Res.* **76**, 2022–2034 (1971).
8. O'Connell, R. J. & Budiansky, B. Seismic velocities in dry and saturated cracked solids. *J. Geophys. Res.* **79**, 5412–5426 (1974).
9. De Fazio, T. L., Aki, K. & Alba, J. Solid earth tide and observed change in the in situ seismic velocity. *J. Geophys. Res.* **78**, 1319–1322 (1973).
10. Reasenber, P. & Aki, K. A precise, continuous measurement of seismic velocity for monitoring in situ stress. *J. Geophys. Res.* **79**, 399–406 (1974).
11. Leary, P. C., Malin, P. E., Phinny, R. A., Brocher, T. & Voncolln, R. Systematic monitoring of millisecond traveltimes variations near Palmdale, California. *J. Geophys. Res.* **84**, 659–666 (1979).
12. Yamamura, K. *et al.* Long-term observation of in situ seismic velocity and attenuation. *J. Geophys. Res.* **108**, doi:10.1029/2002JB002005 (2003).
13. Silver, P. G., Daley, T. M., Niu, F. & Majer, E. L. Active source monitoring of cross-well seismic traveltimes for stress-induced changes. *Bull. Seismol. Soc. Am.* **97**, 281–293 (2007).
14. Aki, K. & Richards, P. G. *Quantitative Seismology* (Freeman, New York, 1980).
15. Abercrombie, R. E. Earthquake source scaling relationships from -1 to $5 M_L$ using seismograms recorded at 2.5-km depth. *J. Geophys. Res.* **100**, 24015–24036 (1995).
16. Rubin, A. M. & Gillard, D. Aftershock asymmetry/rupture directivity among central San Andreas fault microearthquakes. *J. Geophys. Res.* **105**, 19095–19109 (2000).
17. Segall, P., Jonsson, S. & Agustsson, K. When is the strain in the meter the same as the strain in the rock? *Geophys. Res. Lett.* **30**, doi:10.1029/2003GL017995 (2003).
18. Unsworth, M., Bedrosian, P., Eisel, M., Egbert, G. & Siripunvaraporn, W. Along strike variations in the electrical structure of the San Andreas Fault at Parkfield, California. *Geophys. Res. Lett.* **27**, 3021–3024 (2000).

Acknowledgements We thank the NSF funded SAFOD programme and those involved in providing the experiment site, R. Trautz for supplying the barometric pressure logger, M. Zumberge for providing the SAFOD strainmeter data, and D. Lippert and R. Haught for helping with field work. This work was supported by the NSF, Rice University, the Carnegie Institution of Washington, and Lawrence Berkeley National Laboratory of the US Department of Energy under contract DE-AC02-05CH11231.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to F.N. (niu@rice.edu).

METHODS

Data acquisition system. Our acquisition was conducted with a combination of commercial and specially built equipment. The latter are the piezoelectric source and the high voltage amplifier used to power it. The source includes 18 cylindrical rings of piezoelectric ceramic (lead zirconate titanate) epoxied together and wired for positive and negative voltage on the inner and outer surfaces. The source was fluid coupled to the well casing. A three-component accelerometer was clamped to the well casing to provide coupling and reduce relative motions between the source and receiver. We used a commercial recording system, a 'Geode' manufactured by Geometrics, which has a 24 bit analogue-to-digital converter. An air conditioner and heater were used to maintain the recording system electronics within a temperature range of about $\pm 1^\circ\text{C}$.

Triggering was used in our data recording system. The digitizer continually samples the data, and receives a trigger that will generally be between two digitized samples. Including a section of pre-trigger data, the time series is interpolated and re-sampled, so that the new time series begins at the time of the trigger. This start time is not exact, and, at a sampling rate of $48,000\text{ s}^{-1}$, this time is computed to the nearest twentieth of a sample (Geometrics engineering, personal communication). Thus there is a delay time measurement error that will be at most a fortieth of a sample (half-way between samples), and the average error will be an eightieth of a sample, assuming that the errors are uniformly distributed. This corresponds to an average error of 260 ns per trigger. The error in the stacked data decreases by a factor of $N^{1/2}$, assuming the errors are uncorrelated. For $N = 100$, we obtain a timing error of 26 ns.

Optimum experimental design. As shown by Silver *et al.*¹³, there is an optimum distance between the source and receiver that minimizes the detectable threshold of subtle velocity changes:

$$N = Q/\pi \quad (1)$$

Here N is the number of wavelengths between the source and receiver and Q is the quality factor. At the SAFOD site, Q is around 200, which gives $N = 64$. If we assume the S-wave velocity to be 2.8 m ms^{-1} , then the wavelength of the signal with a dominant frequency of 2 kHz is about 1.4 m, so the optimum distance is $\sim 90\text{ m}$. As it was necessary to perform the experiment in the available boreholes, our cross-hole distance was limited to 10 m, which while not optimal still provided us with a good SNR.

Subsample delay time estimate (DTE). In this study, we employed a cosine fitting method to estimate subsample delay time in the time domain^{19,20}. Given the largest sample of the correlation function, $cc(0)$, and its two neighbours $cc(-1)$ and $cc(1)$, the estimated subsample shift is given by following expression:

$$\tau = \alpha / \arctan \left(\frac{cc(-1) - cc(1)}{2cc(0) \sin \alpha} \right) \quad (2)$$

where:

$$\alpha = \arccos \left(\frac{cc(-1) + cc(1)}{2cc(0)} \right) \quad (3)$$

Error estimation. Silver *et al.*¹³ derived a low bound of the error in delay time measurements:

$$\sigma_{\text{DTE}} \geq \frac{1}{2\pi f_0 \text{SNR}} \quad (4)$$

Here f_0 is the dominant frequency of the source pulse, and SNR is the signal-to-noise ratio. Equation (4) indicates that SNR is the only parameter that controls the precision in our delay time estimation when the digitizing error is much less than the background noise in this regime. The precision is not controlled by the sampling rate of the digitizer so it is possible to obtain subsample-interval measurements of the delay time. The dominant frequency of our data is 2 kHz and the SNR is around 700 for the 45-min stacked data, resulting in a best achievable precision of $\sim 1.1 \times 10^{-7}\text{ s}$, or 110 ns in the DTE.

We also measured delay time between each two consecutive samples, which follows a gaussian distribution with a standard deviation of $\sim 80\text{ ns}$ and $\sim 50\text{ ns}$ for the first and second recording period, respectively. In general they are comparable to or even better than the theoretical low bound in equation (4). Since there is contribution from the actual stress-induced velocity perturbations in the measurement, our actual precision can be better than the measured standard deviations. Thus the lower bound appears to be larger than the true DTE error. One possible explanation is that the SNR is significantly underestimated, as the noise is estimated from a time window before the first arrival, which actually contains a considerable amount of non-random electronic noise known as cross-talk, and non-random 'wrap-around' noise from the previous shot.

The precision discussed here does not include other systematic non-random noise, such as changes in the source pulses, errors in trigger timing and digitizer's clock. Such systematic errors could lead to a long-term trend in DTE. To estimate these effects, we also recorded the source pulse waveform in addition to the data. We employed the same method to measure the variation in the source pulse width. Changes in the source pulse width are between $\pm 20\text{ ns}$. This indicates that our source pulse generator and recording system were very stable in the two periods and timing error in the digitizer clock was also very small.

19. Cespedes, I., Huang, Y., Ophir, J. & Spratt, S. Methods for estimation of subsample time delays of digitized echo signals. *Ultrason. Imaging* 17, 142–171 (1995).
20. De Jong, P. G. M., Arts, T., Hoeks, A. P. G. & Reneman, R. S. Determination of tissue motion velocity by correlation interpolation of pulsed ultrasonic echo signals. *Ultrason. Imaging* 12, 84–98 (1990).

The evolutionary origin of flatfish asymmetry

Matt Friedman^{1,2}

All adult flatfishes (Pleuronectiformes), including the gastronomically familiar plaice, sole, turbot and halibut, have highly asymmetrical skulls, with both eyes placed on one side of the head. This arrangement, one of the most extraordinary anatomical specializations among vertebrates, arises through migration of one eye during late larval development. Although the transformation of symmetrical larvae into asymmetrical juveniles is well documented^{1–7}, the evolutionary origins of flatfish asymmetry are uncertain^{1,2} because there are no transitional forms linking flatfishes with their symmetrical relatives^{8,9}. The supposed inviability of such intermediates gave pleuronectiforms a prominent role in evolutionary debates^{10–16}, leading to attacks on natural selection¹¹ and arguments for saltatory change^{14,15}. Here I show that *Amphistium* and the new genus *Heteronectes*, both extinct spiny-finned fishes from the Eocene epoch of Europe, are the most primitive pleuronectiforms known. The orbital region of the skull in both taxa is strongly asymmetrical, as in living flatfishes, but these genera retain many primitive characters unknown in extant forms. Most remarkably, orbital migration was incomplete in *Amphistium* and *Heteronectes*, with eyes remaining on opposite sides of the head in post-metamorphic individuals. This condition is intermediate between that in living pleuronectiforms and the arrangement found in other fishes. *Amphistium* and *Heteronectes* indicate that the evolution of the profound cranial asymmetry of extant flatfishes was gradual in nature.

Teleostei Müller, 1846

Acanthomorpha Rosen, 1973

Percomorpha Rosen, 1973

Pleuronectiformes Bleeker, 1859

Heteronectes chaneti gen. et sp. nov.

Etymology. The generic name references a disused collective name for flatfishes (Heterosomata), as well as the incomplete orbital migration characterizing this taxon (Greek *heteros*, different; *nectri*, swimmer). The specific name honours B. Chanet and his contributions to the study of fossil pleuronectiforms^{8,9,17}.

Holotype. NHMW 1974.1639.24 + 1974.1639.25 (part and counterpart), Naturhistorisches Museum, Vienna (NHMW). Total length is 142 mm; standard length is 111 mm.

Horizon and locality. Bolca, possibly Monte Postale locality, northern Italy. Lower Eocene (Ypresian; SBZ11)¹⁸.

Diagnosis. Stem pleuronectiform differing from *Amphistium* in the following characters: the dorsal- and anal-fin spines are robust, being much broader anteroposteriorly than rays in each fin (rather than a similar thickness); the ventral margin of the anal fin is concave (rather than convex; possible autapomorphy); the anal fin comprises three spines and no fewer than 32 soft, bifurcating rays (rather than 22–21).

The most notable feature of *Heteronectes* is its high degree of cranial asymmetry, manifested primarily in the orbital region of the braincase. This pattern does not arise from taphonomic distortion.

There is neither torsion-induced damage between the orbital region and the posterior of the braincase nor deformation of other components of the skull (Fig. 1a, b).

As in living flatfishes^{2,19}, the frontal bones of *Heteronectes* are unequal in size and shape. The right-side frontal bone is broad, rectangular and ventrally extensive, whereas the left-side frontal bone is reduced to a narrow, curved splint that defines the upper margin of the migrated orbit. The orbital margin fails to extend beyond the dorsal midline. This differs from the arrangement in all living adult flatfishes, in which the migrating eye passes completely over the top of the head to a position on the opposite side of the skull (Fig. 2c). Thus, whereas *Heteronectes* shares a displaced orbit with pleuronectiforms, this genus shows a remarkable intermediate condition in which the eyes remain on opposite sides of the head.

Heteronectes closely resembles *Amphistium*, another percomorph also known from Bolca (*A. paradoxum*)^{9,20} and younger (Lutetian) deposits of France (Paris Basin; *A. altum*)²¹. *Amphistium* has been allied with numerous acanthomorph groups⁹, although many authors noted—and dismissed—a resemblance to pleuronectiforms. The most recent examination of *Amphistium* rejected similarities shared by this genus and flatfishes as primitive or homoplastic, but could not determine its relationships within Percomorpha⁹. Although their systematic interpretations diverged, all previous accounts of *Amphistium* agreed that this genus has a symmetrical skull. Renewed study, using conventional techniques and computed tomography, unequivocally shows that *Amphistium* is characterized by conspicuous cranial asymmetry centred in the orbital region (Figs 1c–f and 2b). Just as in *Heteronectes*, the migrated orbit of *Amphistium* does not extend beyond the dorsal midline. The mode and degree of asymmetry is identical between multiple specimens of *Amphistium*, clearly demonstrating that it reflects a genuine anatomical pattern rather than post-mortem deformation.

Many lines of evidence indicate that the observed cranial morphology in these Eocene fossils represents the adult condition, not an ontogenetic stage preceding complete orbital migration characteristic of all other flatfishes. First, all individuals of *Amphistium* and *Heteronectes* are an order of magnitude longer than the size at which eye migration begins in most living flatfishes (between 10 mm and 15 mm)³. There is no indication that primitive flatfishes might have metamorphosed at larger body sizes, because the eye passes over the dorsal midline at 13 mm in the anatomically conservative *Psettodes*²². Second, unlike living flatfishes in even the late stages of eye migration^{4,5,7}, *Amphistium* and *Heteronectes* have completely mineralized skulls, with all component ossifications present. Third, the consistent morphology of all *Amphistium* specimens—which show identical degrees of asymmetry despite ranging in size between 103 mm (MNHN 10878b/Bol87) and 200 mm (BMNH P.16138) in standard length—indicates this is the adult arrangement.

Amphistium and *Heteronectes* share with flatfishes the specialization of displaced orbits, an unambiguous synapomorphy of that clade¹⁹. Several other derived features, unrelated to asymmetry, link

¹Committee on Evolutionary Biology, University of Chicago, 1025 East 57th Street, Chicago, Illinois 60637, USA. ²Department of Geology, The Field Museum, 1400 South Lake Shore Drive, Chicago, Illinois 60605, USA.

Amphistium to flatfishes. These characters, which cannot be checked in the incomplete material of *Heteronectes*, include an anteriorly extensive first dorsal fin, a procumbent first pterygiophore of the dorsal fin, and anteriorly curved neural spines in the abdominal region. However, apart from these flatfish synapomorphies, *Amphistium* and *Heteronectes* display a series of generalized percomorph characters lost or transformed in other pleuronectiforms. These genera retain a single supraneural (uncertain in *Heteronectes*; absent in crown-group pleuronectiforms), lack a pseudomesial bar (a neomorph that frames the migrated orbit in crown-group pleuronectiforms⁶), and possess a generalized 'perciform' caudal endoskeleton²³ comprising a full complement of uroneurals, epurals and independent hypurals. *Amphistium* and *Heteronectes* also bear fin spines in both the anal and dorsal fins. This diagnostic acanthomorph

feature^{24,25} is only found in the plesiomorphic *Psettodes* among living flatfishes^{2,19}. In the context of a phylogenetic analysis, this distribution of characters indicates that *Amphistium* and *Heteronectes* branch from pleuronectiform stem, outside the crown clade that contains living forms plus all other known fossil examples (Fig. 2a). Placement of *Heteronectes* and *Amphistium* as successive taxa along the flatfish stem is notable because it suggests that the observed asymmetry in these genera represents an evolutionary intermediate between generalized conditions and the arrangement found in crown-group pleuronectiforms, rather than a secondary development of incomplete orbital migration from an ancestor showing complete transit.

Specimens of *Amphistium* include both dextral and sinistral morphs (Figs 1c–f and 2b). This is consistent with the pattern seen

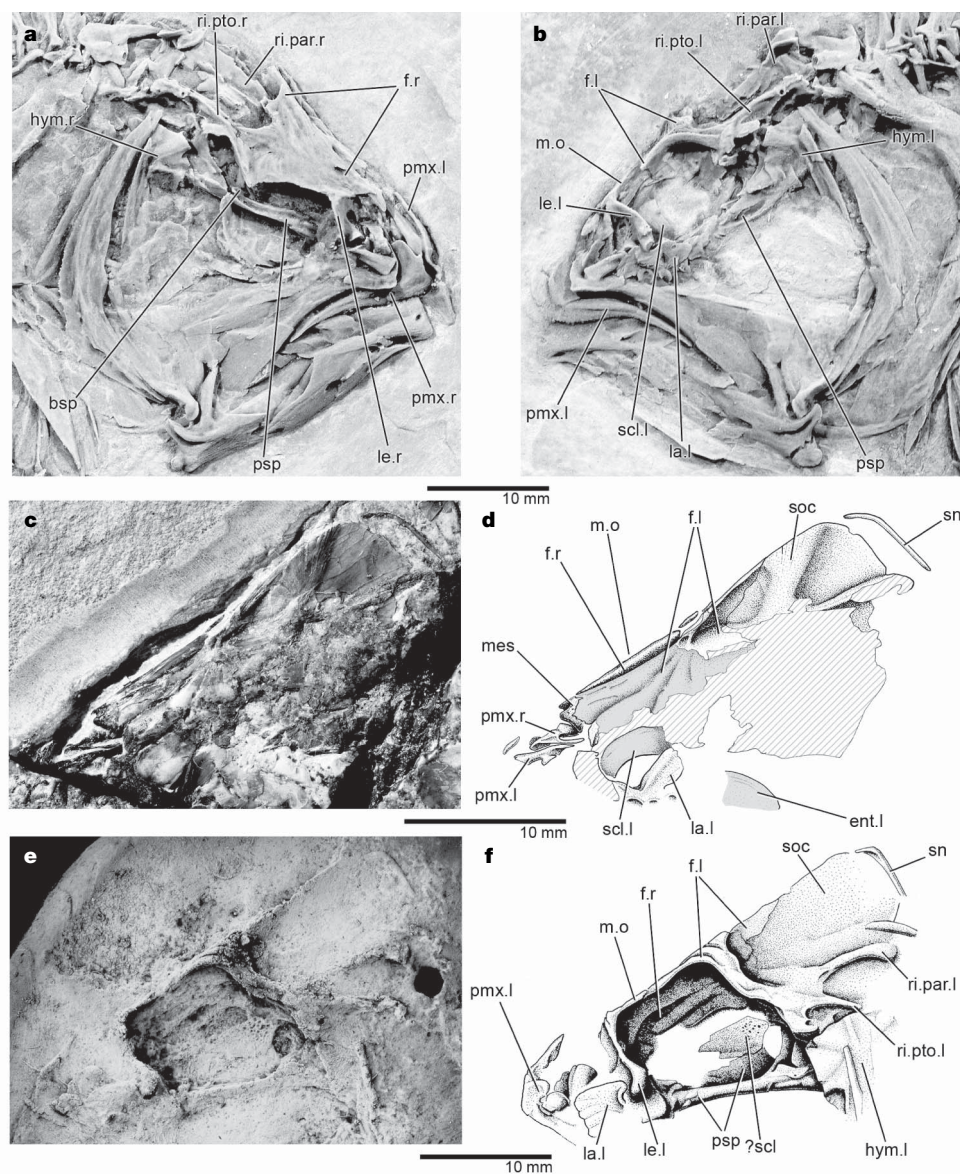


Figure 1 | Skulls of primitive pleuronectiforms showing incomplete orbital migration intermediate between generalized fishes and living flatfishes.

a, *Heteronectes chaneti* gen. et sp. nov., holotype, NHMW 1974.1639.25 (dextral morph); transfer preparation dusted with ammonium chloride and presented in right-lateral view. **b**, Counterpart, NHMW 1974.1639.24; transfer preparation dusted with ammonium chloride and presented in left-lateral view, showing migrated orbit. **c**, *Amphistium paradoxum*, Muséum national d'Histoire naturelle, Paris (MNHN), MNHN 10878b/Bol87 (sinistral morph); specimen presented in left-lateral view (photo credit: C. Lemzaouda, MNHN). **d**, Interpretive drawing. Solid grey shading indicates

impression; diagonal hatching indicates damaged bone. **e**, *Amphistium altum*, Natural History Museum, London (BMNH), BMNH P. 3940 (dextral morph); silicone peel dusted with ammonium chloride and presented in left-lateral view, showing migrated orbit. **f**, Interpretive drawing. bsp, basisphenoid; ent, entopterygoid; f, frontal; hym, hyomandibular; la, lacrimal; le, lateral ethmoid; m.o, dorsal margin of migrated orbit; mes, mesethmoid; pmx, premaxilla; psp, parasphenoid; ri.par, parietal/epioccipital ridge; ri.pto, pterotic ridge; scl, sclerotic ring; sn, supraneural; soc, supraoccipital; names followed by 'r' or 'l' indicate right or left feature, respectively; '?' indicates uncertain identification. Scale bars, 10 mm.

in the morphologically primitive living flatfish *Psettodes*, where the two morphs occur in near equal frequency, but differs from the condition in more derived forms, which typically occur almost exclusively as one morph or the other²⁶. *Amphistium* indicates that indiscriminate orbital migration is primitive for flatfishes, with dominant left- or right-eyed migration representing a derived arrangement. This conforms to a general pattern in the evolution of directional asymmetries that arise late in development, in which groups often first pass through an antisymmetric stage where individuals occur as both dextral and sinistral forms at near-equal frequencies²⁷.

Amphistium and *Heteronectes* deliver the first clear picture of flatfish origins, a hotly contested issue in debates on the mode and tempo of evolution owing to the unclear adaptive value of incomplete eye migration. Flatfishes formed a cornerstone of early arguments against natural selection¹¹, forcing Darwin¹² to respond with a scenario that invoked the inheritance of acquired traits, similar to Lamarck's theory of flatfish origins espoused 60 years earlier¹⁰. Goldschmidt asserted that the first flatfish must have arisen suddenly as a "hopeful monster", bypassing any intermediate forms, and gave this group a

key position in his arguments for saltatory evolution^{14,15}. The discovery of stem flatfishes with incomplete orbital migration refutes these claims and demonstrates that the assembly of the flatfish bodyplan occurred in a gradual, stepwise fashion. Thus, the evolutionary origin of flatfish asymmetry resembles its developmental origin, with increasing degrees of orbital migration transforming a symmetrical precursor into a fully asymmetrical form (Fig. 2c).

Questions about the possible selective advantage of incomplete orbital transit arise from the discovery of stem flatfishes. Clues are given by living taxa, which often prop their bodies above the substrate by depressing their dorsal- and anal-fin rays²⁸. Similar behaviour might have permitted *Amphistium* and *Heteronectes*—both of which have long median-fin rays—the use of both eyes while on the sea floor. The unusual morphology and resting orientation of pleuronectiforms have been interpreted as adaptations for prey ambush¹⁶, and it is clear that stem flatfishes, like morphologically primitive living forms, were piscivorous; one specimen of *Amphistium* (MCSNV V.D.91+92) contains the remains of a fish nearly half its own length. The fossil record rejects the intuitive notion that

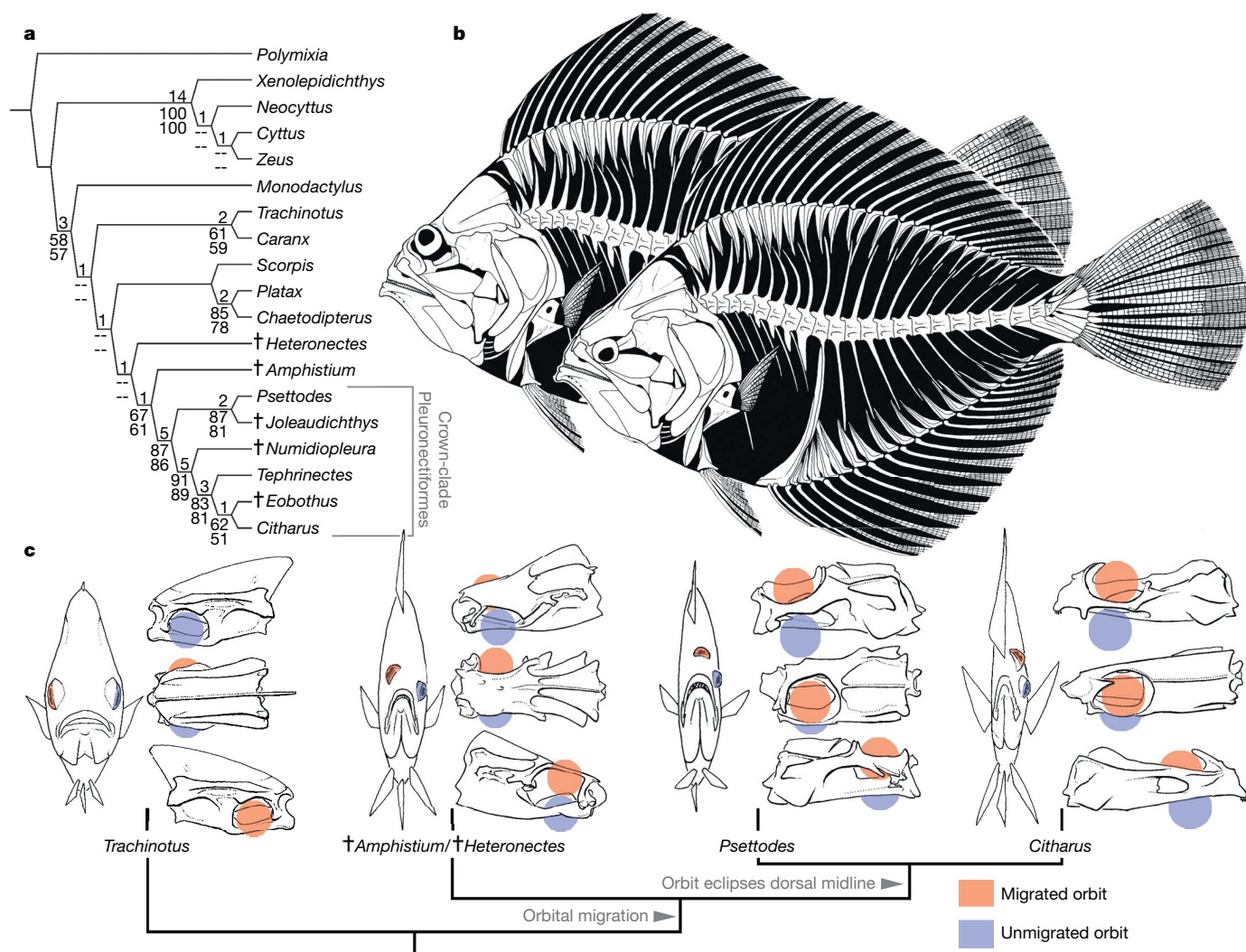


Figure 2 | Phylogenetic placement of *Heteronectes* and *Amphistium* and implications for the origin of cranial asymmetry in flatfishes. **a**, Topology arising from the analysis of a matrix comprising 19 taxa coded for 58 morphological characters (8 ordered) (number of cladograms = 1; cladogram length = 135; consistency index = 0.50; retention index = 0.74; rescaled consistency index = 0.37). *Heteronectes* and *Amphistium* are placed as successively more crownward plesions on the flatfish stem. Unordered analyses reconstruct these taxa in the same position. Numbers at nodes indicate Bremer decay index, bootstrap support and jackknife support, from

top to bottom, respectively. Extinct taxa are marked (†) and '–' indicates that bipartition occurs in fewer than half of cladograms arising from bootstrap or jackknife analysis. Previous placements of *Amphistium* outside Pleuronectiformes are rejected. See Supplementary Information for full details of the analysis. **b**, Reconstruction of *Amphistium*, showing sinistral (front) and dextral (back) individuals in the left lateral view (modified from ref. 20). **c**, Simplified cladogram adapted from **a** showing the progression of orbital migration across flatfish phylogeny. Neurocrania are depicted in left lateral (top), dorsal (middle) and right lateral (bottom) views.

incomplete orbital migration might be maladaptive. Stem flatfishes with this condition range over two geological stages and derive from localities that also yield crown pleuronectiforms with full cranial asymmetry^{8,9}.

Amphistium and *Heteronectes* are contemporaries of the earliest members of many derived pleuronectiform lineages^{8,29}, including the oldest known sole¹⁷. The sudden appearance of anatomically modern pleuronectiform groups in the Palaeogene period matches the pattern repeated by many acanthomorph clades^{24,29}. Inferring interrelationships between higher groups in this explosive radiation has proved difficult, and an unresolved bush persists³⁰. Documenting the origin of these clades is vital to understanding the roots of modern biodiversity, because acanthomorph fishes comprise nearly one-third of living vertebrate species³⁰. Stem representatives—such as *Amphistium* and *Heteronectes* in the case of pleuronectiforms—have yet to be identified for many acanthomorph clades^{24,29}, but their recognition might prove invaluable in delivering a stable hypothesis of interrelationships for this exceptional vertebrate radiation.

METHODS SUMMARY

Computed tomography scans were carried out the University of Texas at Austin. The positions of *Amphistium* and *Heteronectes* were inferred by parsimony analysis of a morphological data set. Details of these analyses and further anatomical data are given in Supplementary Information.

Received 8 February; accepted 22 May 2008.

- Kyle, H. M. The asymmetry, metamorphosis, and origin of flat-fishes. *Phil. Trans. R. Soc. Lond. B* **211**, 75–129 (1923).
- Norman, J. R. *A Systematic Monograph of the Flatfishes (Heterosomata). Vol. 1. Psettodidae, Bothidae, Pleuronectidae* (British Museum (Natural History), London, 1934).
- Ahlstrom, E. H., Amaoka, K., Hensley, D. A., Moser, H. G. & Sumida, B. Y. in *Ontogeny and Systematics of Fishes* (eds Moser, H. G. et al.) 640–670 (Allen, Lawrence, Kansas, 1984).
- Okada, N., Takagi, Y., Seikai, T., Tanaka, M. & Tagawa, M. Asymmetrical development of bones and soft tissues during eye migration of metamorphosing Japanese flounder, *Paralichthys olivaceus*. *Cell Tissue Res.* **304**, 59–66 (2001).
- Wagemans, F. & Vandewalle, P. Development of the bony skull in common sole: brief survey of morpho-functional aspects of ossification sequence. *J. Fish Biol.* **59**, 1350–1369 (2001).
- Hoshino, K. Fixing the confused term “pseudomesial bar” and homologies of pleuronectiform cranial elements, with proposals of new terms. *Ichthyol. Res.* **53**, 435–440 (2006).
- Schreiber, A. M. Asymmetric craniofacial remodeling and lateralized behavior in larval flatfish. *J. Exp. Biol.* **209**, 610–621 (2006).
- Chanet, B. A cladistic reappraisal of the fossil flatfishes record consequences on the phylogeny of the Pleuronectiformes (Osteichthyes: Teleostei). *Ann. Sci. Nat. Zool.* **18**, 105–117 (1997).
- Chanet, B. Supposed and true flatfishes [Teleostei: Pleuronectiformes] from the Eocene of Monte Bolca, Italy. *Stud. Ric. Giacim. Terz. Bolca* **8**, 220–243 (1999).
- Lamarck, J. B. *Philosophie Zoologique* (Dentu, Paris, 1809).
- Mivart, St G. J. *On the Genesis of Species* (MacMillan, London, 1871).
- Darwin, C. *On the Origin of Species by Means of Natural Selection* 6th edn (Murray, London, 1872).
- Wallace, A. R. *Darwinism* (MacMillan and Co., London, 1889).
- Goldschmidt, R. Some aspects of evolution. *Science* **78**, 539–547 (1933).
- Goldschmidt, R. *The Material Basis of Evolution* (Yale Univ. Press, New Haven, 1940).
- von Wahlert, G. The role of ecological factors in the origin of higher levels of organization. *Syst. Zool.* **14**, 288–300 (1965).
- Chanet, B. *Eubuglossus eocenicus* (Woodward 1910) from the Upper Lutetian of Egypt, one of the oldest soleids (Teleostei, Pleuronectiformes). *Neues Jahrb. Geol. Paläontol. Monatsh.* **1994**, 391–398 (1994).
- Trevisani, E., Papazzoni, C. A., Ragazzi, E. & Roghi, G. Early Eocene amber from the “Pesciara di Bolca” (Lessini Mountains, Northern Italy). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **223**, 260–274 (2005).
- Chapleau, F. Pleuronectiform relationships: a cladistic reassessment. *Bull. Mar. Sci.* **52**, 516–540 (1993).
- Blot, J. Les poissons fossiles du Monte Bolca. Classés jusqu’ici dans les familles des Carangidae, Menidae, Ephippidae, Scatophagidae. *Mem. Mus. Civ. Stor. Nat. Verona* **1**, 1–525 (1969).
- Gaudant, J. Sur un *Amphistium* (poisson téléostéen, Percoidei) de l’Éocène Parisien. *Bull. Inf. Géol. Bass. Paris* **16**, 11–14 (1979).
- Liew, H.-C., Milward, N. E. & Hartwick, R. F. Descriptions of larval fishes of the genera *Psettodes* (Psettodidae) and *Pseuorhombus* (Paralichthyidae) from the Great Barrier Reef, Australia. *Aust. J. Mar. Freshwat. Res.* **39**, 51–70 (1988).
- Gosline, W. A. The perciform caudal skeleton. *Copeia* **1961**, 265–270 (1961).
- Patterson, C. An overview of the early fossil record of acanthomorphs. *Bull. Mar. Sci.* **52**, 29–59 (1993).
- Johnson, G. D. & Patterson, C. Percomorph phylogeny: a survey of acanthomorphs and a new proposal. *Bull. Mar. Sci.* **52**, 554–626 (1993).
- Hubbs, C. L. & Hubbs, L. C. Bilateral asymmetry and bilateral variation in fishes. *Pap. Mich. Acad. Sci. Arts Lett.* **30**, 229–310 (1944).
- Palmer, A. R. From symmetry to asymmetry: phylogenetic patterns of asymmetry variation in animals and their evolutionary significance. *Proc. Natl Acad. Sci. USA* **93**, 14279–14286 (1996).
- Stickney, R. R., White, D. B. & Miller, D. Observations of fin use in relation to feeding and resting behavior in flatfishes (Pleuronectiformes). *Copeia* **1973**, 154–156 (1973).
- Patterson, C. in *The Fossil Record 2* (ed. Benton, M. J.) 621–656 (Chapman & Hall, London, 1993).
- Stiassny, M. L. J., Wiley, E. O., Johnson, G. D. & de Carvalho, M. R. in *Assembling the Tree of Life* (eds Cracraft, J. & Donoghue, M. J.) 410–429 (Oxford Univ. Press, New York, 2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements I thank G. Clément, P. Forey, D. Goujet, M. Richter, O. Schultz, A. Vaccari and M. Vêran for loaning or providing access to fossil specimens; R. Arrindell, B. Brown, D. Johnson, S. Raredon, M. Rogers and M. Westneat for arranging the loan or study of recent material; E. Hilton for checking gill-arch characters in carangids; M. Colbert and A. Gosselin-Ildari for compiling the computed tomography renderings, K. Claeson for providing specimen transport; L. Herzog, A. Shinya, D. Wagner and J. Holstein for helping in fossil preparation; and A. Bannikov, M. Coates, M. LaBarbera and N. Smith for discussion. This work was supported by a grant from the Lerner-Grey Fund for Marine Research, a Hinds Fund Grant, an Evolving Earth Grant, a National Science Foundation Graduate Research Fellowship (award number DGE-0228235), and an Environmental Protection Agency STAR Fellowship (award number FP916730).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.F. (mattf@uchicago.edu).

Social diversity promotes the emergence of cooperation in public goods games

Francisco C. Santos¹, Marta D. Santos² & Jorge M. Pacheco²

Humans often cooperate in public goods games^{1–3} and situations ranging from family issues to global warming^{4,5}. However, evolutionary game theory predicts^{4,6} that the temptation to forgo the public good mostly wins over collective cooperative action, and this is often also seen in economic experiments⁷. Here we show how social diversity provides an escape from this apparent paradox. Up to now, individuals have been treated as equivalent in all respects^{4,8}, in sharp contrast with real-life situations, where diversity is ubiquitous. We introduce social diversity by means of heterogeneous graphs and show that cooperation is promoted by the diversity associated with the number and size of the public goods game in which each individual participates and with the individual contribution to each such game. When social ties follow a scale-free distribution⁹, cooperation is enhanced whenever all individuals are expected to contribute a fixed amount irrespective of the plethora of public goods games in which they engage. Our results may help to explain the emergence of cooperation in the absence of mechanisms based on individual reputation and punishment^{10–12}. Combining social diversity with reputation and punishment will provide instrumental clues on the self-organization of social communities and their economical implications.

The N -person prisoner's dilemma constitutes the most used metaphor to study public goods games (PGGs): cooperators (C) contribute an amount c ('cost') to the public good; defectors (D) do not contribute. The total contribution is multiplied by an enhancement

factor r and the result is equally distributed between all N members of the group. Hence, Ds get the same benefit of the Cs at no cost. Collective action to shelter, protect and nourish, which abounds in the animal world, provides examples of PGGs, because the cooperation of group members is required. Ultimately, the success (and survival)⁵ of the human species relies on the capacity of humans for large-scale cooperation. In the absence of enforcement mechanisms^{7,13,14}, conventional evolutionary game theory predicts that the temptation to defect leads individuals to forgo the public good⁴ in the N -person prisoner's dilemma. Whenever interactions are not repeated, and reward and punishment^{4,8,13} can be ruled out, several mechanisms were explored that promote cooperation. Individuals were either constrained to interact only with their neighbours on spatial lattices^{4,8,15}, or given the freedom to opt out of participating^{4,15}, leading to a coexistence of cooperators and defectors, even on spatial lattices.

Here we investigate what happens in the absence of reputation and punishment and when participation is compulsory. Unlike previous studies involving PGGs^{4,8,15}, individuals now interact along the social ties defined by a heterogeneous graph^{16–18}. This reflects the fact that individuals have different roles in social communities. Empirical studies show that social graphs have a marked degree of heterogeneity combined with small-world effects^{19–21}, as illustrated in Fig. 1b. Hence we introduce diversity in the study of cooperation under PGGs in the context of evolutionary graph theory²², and in the presence of spatial and network reciprocity^{18,23,24}.

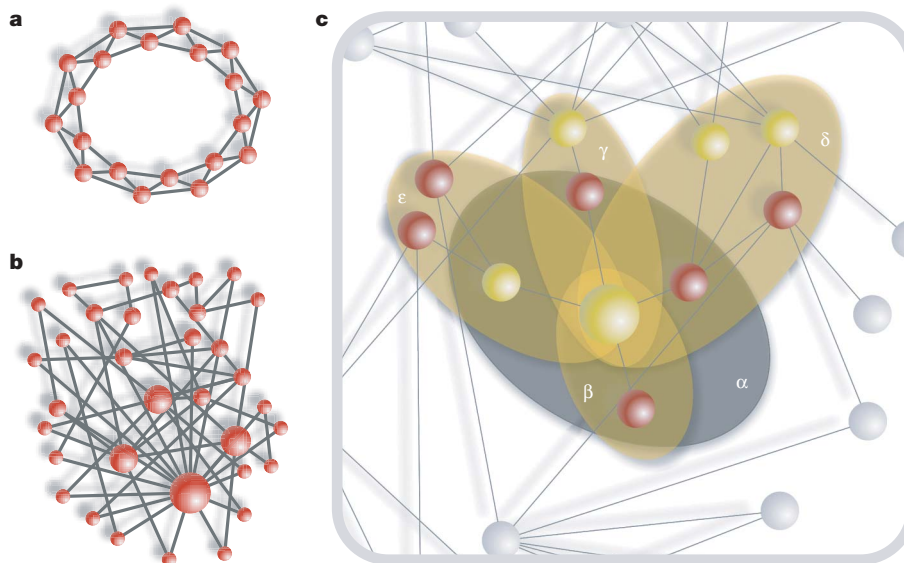


Figure 1 | Population structure and local neighbourhoods. **a**, Regular graphs studied so far, which mimic spatially extended systems. **b**, Scale-free graphs⁹ in which small-world effects coexist with a large heterogeneity in neighbourhood size. **c**, The focal individual (largest sphere) belongs to different groups (neighbourhoods) of different sizes in a heterogeneous graph. Given his/her connectivity $k = 4$, we identify five neighbourhoods, each centred on one of the members of the focal individual's group, such that individual fitness derives from the payoff accumulated in all five neighbourhoods (α , β , γ , δ and ϵ).

¹Institut de Recherches Interdisciplinaires et de Développements en Intelligence Artificielle (IRIDIA), Computer and Decision Engineering Department, Université Libre de Bruxelles, B-1050 Brussels, Belgium. ²ATP-group, Centro de Física Teórica e Computacional (CFTC) and Departamento de Física da Universidade de Lisboa, Complexo Interdisciplinar, Av. Prof. Gama Pinto 2, 1649-003 Lisboa, Portugal.

Figure 1c shows how diversity is introduced, in which we enumerate the different PGGs in which the focal individual (large sphere) engages. Each PGG is associated with a fixed neighbourhood defined by the social graph; given the focal individual's connectivity $k = 4$, he/she participates in five PGGs; that centred in his/her neighbourhood α (group size of 5) plus those associated with the neighbourhoods centred on his/her neighbours: β (2), γ (3), δ (5) and ϵ (4). Hence graph heterogeneity leads individuals to engage in different numbers of PGGs with different group sizes. Furthermore, there is no reason for every C to contribute the same amount to each game in which he/she participates (see below).

Figure 2a shows results for the evolution of cooperation corresponding to the conventional situation in which every C pays a fixed cost c in every game that he/she plays. We plot the fraction of cooperators in the population that survive evolution as a function of the renormalized PGG enhancement factor $\eta = r/(z + 1)$, where z is the average connectivity of the population graph (see Methods). In infinite, well-mixed populations, a sharp transition from defection to cooperation takes place at $\eta = 1$. Comparison between the results obtained on regular graphs (Fig. 1a) with those on strongly heterogeneous graphs (scale-free; Fig. 1b) reveal the sizable impact of heterogeneity on the evolution of cooperation. For regular graphs (in which, from the perspective of a population structure, every individual is equivalent to any other) cooperators become predominant (their fraction exceeds 50%) at $\eta \approx 0.7$: network reciprocity^{18,23,24} leads to an enhancement of cooperation also under PGGs^{4,8,15}. This number decreases to $\eta \approx 0.6$ on scale-free graphs, in which individual participation now reflects both a diversity in

the size of each individual's PGGs and in the different number of PGGs in which each individual participates.

The contribution of each C in Fig. 2a has been proportional to $k + 1$, where k is the number of neighbours (vertex degree). This may be unrealistic, because individuals have limited resources and social rules often accommodate a more egalitarian overall contribution from individuals²⁵. In the extreme opposite limit, all Cs contribute the same overall cost, equally shared between all games in which each individual participates. In this limit, still another new type of diversity is introduced—that of individual contributions to each game. Real-world situations will naturally fall somewhere between these limits, as individuals learn²⁶ to cooperate (or defect) in better ways. In general, however, one expects diversity of contributions from individuals. Depending on the problem at stake, any contribution may be necessary and even welcome, however small. Below we show that, whenever all contributions are interpreted as acts of cooperation, cooperation blooms.

Figure 2b shows the results including this additional diversity in which Cs contribute $c/(k + 1)$ for each game, k being their degree in the social graph. This new model leads to an impressive boost of cooperation. In all cases, cooperation now dominates for values of η below 0.4.

What is the origin of such a boost of cooperation? Because each C now contributes $c/(k + 1)$ to each game, diversity resulting from heterogeneous graphs determines a richer spectrum of individual fitness. In a single PGG, the fitness difference between a C and a D is no longer constant and proportional to c , as on homogeneous graphs, but now depends on the social context of the individual. As shown in detail in Supplementary Information, the highly connected nodes (hubs) are those that turn most quickly into cooperation. This is because, under this contribution model, the relative fitness of a single cooperator increases with its connectivity, as illustrated in Fig. 3 (derivation details are given in Supplementary Information). Consequently, heterogeneity confers a natural advantage on hubs. In practice, Cs survive extinction for values of η about 0.25. Because $z = 4$, $\eta = 0.25$ implies that $r = 1.25$, much lower than the size $N = 3$ of the smallest group in the entire population. Note that, in those games for which $\eta_k \equiv r/(k + 1) > 1$ (the smaller groups), the social dilemma might become relaxed, because in this case it is better to play C than D. As Fig. 2b shows, cooperation prevails despite $\eta_k < 1$ in every PGG played. In fact, the impact of diversity is preserved even when the social dilemma is transformed such that defection is always preferred, irrespective of η .

It still remains to explain how a D sitting on a large hub can be taken over by a C. As shown in the Supplementary Information, Ds are victims of their own success—successful Ds breed Ds in their neighbourhood, inducing a negative feedback mechanism that reduces their fitness²⁷. Consequently, they become vulnerable to nearby cooperators. Once invaded by a C, a hub will remain C, as by placing Cs on nearby sites, successful Cs increase their fitness. The role of Cs is therefore crucial and twofold: they efficiently disseminate the cooperator strategy across social networks, whereas they get a stronghold on hubs by minimizing the potential loss from exploitation by free-riding Ds. It is noteworthy that the results shown in Fig. 2b, in which selection is strong, are robust with respect to the detailed evolutionary dynamics (pairwise comparison²⁸, birth-death¹⁸, death-birth¹⁸), to the updating strategy (synchronous, asynchronous), and even to errors (mutations cannot destroy C-dominance).

What about the defectors? They have a minor role as social parasites when they survive on such graphs. Figure 2b shows that some residual Ds continue to exploit Cs. In the Supplementary Information we provide a detailed analysis showing how the evolutionary dynamics inexorably leads cooperators to invade the hubs quickly, whereas defectors are able to survive only on loosely connected nodes, with low fitness and exploiting cooperators of low fitness.

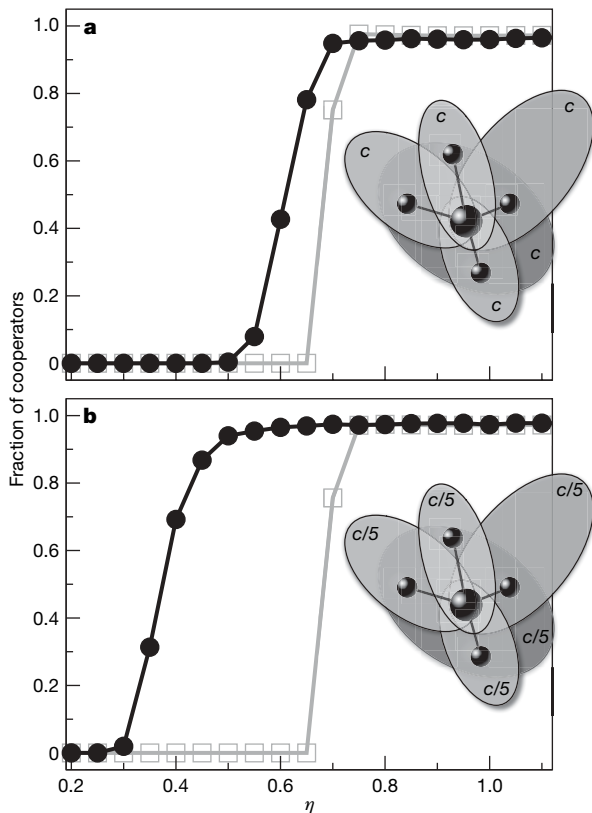


Figure 2 | Evolution of cooperation in networked PGGs. Black lines and filled circles show results for scale-free graphs; grey lines and open squares show results for regular graphs. In all cases $z = 4$. **a**, Fixed cost per game. Cs pay a cost $c = 1$ for each PGG in which they participate: diversity in number of PGGs and size of each PGG associated with scale-free graphs merits a significant enhancement of cooperation. **b**, Fixed cost per individual. Each C contributes a total cost c equally shared between all $k + 1$ PGGs in which he/she engages. This change of model leads to an impressive boost of cooperation.

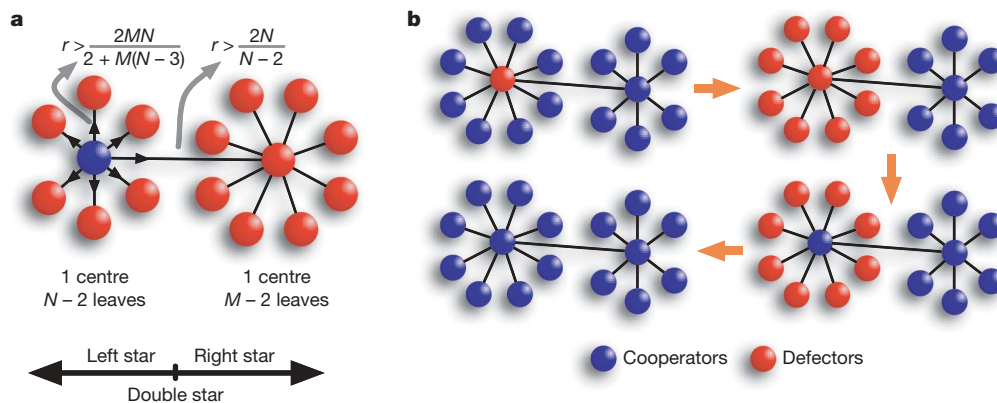


Figure 3 | Dynamics on the double star. **a**, When a single C (blue) occupies the centre of the left star, the critical r above which his/her fitness exceeds that of any of the leaves decreases whenever the size (number of leaves) of the left star increases or the size of the right star decreases. Taking over the D

(red) on the right centre does not depend on his/her connectivity. **b**, When a D occupies a high-fitness location, the fact that he/she places other Ds in his/her neighbourhood leads to his/her own demise (see Supplementary Information for details).

In a more economical perspective, our results also portray different evolutionary outcomes even in communities in which all individuals cooperate. Now we consider populations of 100% cooperators and look at their 'wealth' (fitness) distribution according to different underlying models. We consider homogeneous (regular) and heterogeneous (scale-free) graphs. In Fig. 4 we plot the fraction of the population that holds a given fraction of the total wealth.

The differences are striking. On regular graphs an egalitarian wealth distribution is obtained, irrespective of the contribution model. On scale-free graphs wealth distributions follow a power law. However, for a fixed cost per individual, the population has significantly fewer poor and more rich (note the logarithmic scale in Fig. 4). Given that the emergence of cooperation is easiest in this case, the results provide an impressive account of the role of diversity and its implications in both the emergence of cooperation and the resulting wealth distribution.

In this study any contribution has been identified with cooperation. In communities under the influence of social norms, individual contributions will be easily classified as acts of cooperation (or not). In this context, our results suggest the possibility that successful communities are those in which the act of giving is more important than the amount given. This may be of particular relevance whenever the survival of the community is at stake, in which case any help is

necessary^{14,25}. Most probably, in such cases selection is strong, as considered here.

METHODS SUMMARY

According to Fig. 1c, each individual and his/her k neighbours statically define a group (of size $k + 1$). The fitness of individual i is associated with the accumulated payoff resulting from all PGGs in which he/she participates. Strategy evolution is implemented by using the finite population analogue of the replicator dynamics: at each time step each individual will adopt the strategy of a randomly chosen neighbour (if more fit) with a probability proportional to the fitness difference^{16,29}. Consequently, the results become independent of the specific value used for the cost of cooperation c (we set c to 1). The results shown were obtained for communities of 10^3 individuals starting with 50% of cooperators randomly distributed on the population graph. The equilibrium fraction of cooperators results from averaging over 2,000 generations after a transient period of 10^5 generations. This procedure was repeated 100 times for 10 different realizations of each class of graph. Finally, the distributions depicted in Fig. 4 were obtained by averaging the fitness distributions over 50 scale-free graphs with average connectivity $z = 4$ and populations of 10^3 individuals, all cooperators.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 23 December 2007; accepted 20 March 2008.

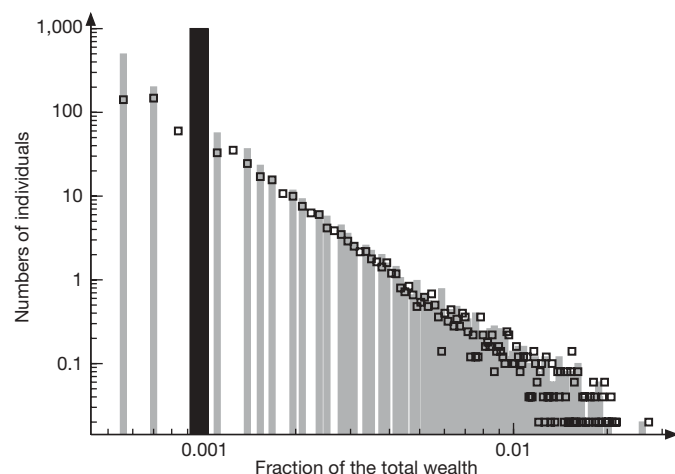


Figure 4 | Wealth distribution. On scale-free graphs, the wealth (fitness) distribution follows a power law, both when each individual invests a fixed cost per game (thin grey bars) or when he/she spends an overall fixed cost (open squares). The latter case leads to communities with fewer 'poor' and more 'rich'. These behaviours contrast with the egalitarian wealth distribution characteristic of homogeneous graphs (thick black bar).

- Hardin, G. The tragedy of the commons. *Science* **162**, 1243–1248 (1968).
- Binmore, K. G. *Game Theory and the Social Contract* Vol. 1, *Playing Fair* (MIT Press, Cambridge, MA, 1994).
- Kollock, P. Social dilemmas: The anatomy of cooperation. *Annu. Rev. Sociol.* **24**, 183–214 (1998).
- Hauert, C., De Monte, S., Hofbauer, J. & Sigmund, K. Volunteering as Red Queen mechanism for cooperation in public goods games. *Science* **296**, 1129–1132 (2002).
- Milinski, M., Semmann, D., Krambeck, H. J. & Marotzke, J. Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proc. Natl Acad. Sci. USA* **103**, 3994–3998 (2006).
- Hofbauer, J. & Sigmund, K. *Evolutionary Games and Population Dynamics* (Cambridge Univ. Press, Cambridge, 1998).
- Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
- Szabo, G. & Hauert, C. Phase transitions and volunteering in spatial public goods games. *Phys. Rev. Lett.* **89**, 118101 (2002).
- Barabási, A. L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
- Ohtsuki, H. & Iwasa, Y. How should we define goodness?—reputation dynamics in indirect reciprocity. *J. Theor. Biol.* **231**, 107–120 (2004).
- Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity. *Nature* **437**, 1291–1298 (2005).
- Pacheco, J. M., Santos, F. C. & Chalub, F. A. Stern-judging: A simple, successful norm which promotes cooperation under indirect reciprocity. *PLoS Comput. Biol.* **2**, e178 (2006).
- Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A. & Sigmund, K. Via freedom to coercion: the emergence of costly punishment. *Science* **316**, 1905–1907 (2007).

14. Boyd, R. & Mathew, S. Behavior. A narrow road to cooperation. *Science* **316**, 1858–1859 (2007).
15. Szabo, G. & Hauert, C. Evolutionary prisoner's dilemma games with voluntary participation. *Phys. Rev. E* **66**, 062903 (2002).
16. Santos, F. C. & Pacheco, J. M. Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.* **95**, 098104 (2005).
17. Santos, F. C., Pacheco, J. M. & Lenaerts, T. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc. Natl Acad. Sci. USA* **103**, 3490–3494 (2006).
18. Ohtsuki, H., Hauert, C., Lieberman, E. & Nowak, M. A. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505 (2006).
19. Amaral, L. A., Scala, A., Barthelemy, M. & Stanley, H. E. Classes of small-world networks. *Proc. Natl Acad. Sci. USA* **97**, 11149–11152 (2000).
20. Albert, R. & Barabási, A. L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
21. Dorogotsev, S. N. & Mendes, J. F. F. *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford Univ. Press, Oxford, 2003).
22. Lieberman, E., Hauert, C. & Nowak, M. A. Evolutionary dynamics on graphs. *Nature* **433**, 312–316 (2005).
23. Nowak, M. A. & May, R. M. Evolutionary games and spatial chaos. *Nature* **359**, 826–829 (1992).
24. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
25. Boehm, C. *Hierarchy in the Forest: The Evolution of Egalitarian Behavior* (Harvard Univ. Press, Cambridge, MA, 1999).
26. Skyrms, B. *The Stag Hunt and the Evolution of Social Structure* (Cambridge Univ. Press, Cambridge, 2004).
27. Santos, F. C. & Pacheco, J. M. A new route to the evolution of cooperation. *J. Evol. Biol.* **19**, 726–733 (2006).
28. Traulsen, A., Nowak, M. A. & Pacheco, J. M. Stochastic dynamics of invasion and fixation. *Phys. Rev. E* **74**, 011909 (2006).
29. Hauert, C. & Doebeli, M. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* **428**, 643–646 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements F.C.S. acknowledges support from Fonds de la Recherche Scientifique, Belgium. M.D.S. and J.M.P. acknowledge financial support from Fundação para a Ciência e Tecnologia, Portugal.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.M.P. (pacheco@cii.fc.ul.pt).

METHODS

Population structure. Population structure is represented by a graph; individuals occupy the vertices of the graph, and social interactions proceed along the edges. Figure 1 depicts the two topologies considered in this study. Figure 1a represents one-dimensional lattices, namely communities in which all individuals (nodes) are topologically equivalent, similar to those previously investigated. In Fig. 1b we provide a diagram of more realistic social structures, portraying populations in which different individuals have a distinct number of connections. Such strongly heterogeneous populations were obtained by means of the Barabási–Albert scale-free model based on growth and preferential attachment⁹. As is well known, realistic social structures fall somewhere between regular and scale-free graphs¹⁹. The blue circles in Fig. 2 were obtained on scale-free populations with 10^3 individuals generated in this way, with average connectivity $z = 4$. This implies that the smallest group in this population has three individuals⁹. In all simulations, the networks remain unchanged throughout evolution and each individual adopts a pure strategy: cooperator (C) or defector (D).

Public goods games. Every individual participates in all possible PGGs, accumulating the benefits and costs resulting from each of them. The accumulated value of the payoffs resulting from all possible PGGs contributes to individual fitness (neighbourhoods α , β , γ , δ and ε in Fig. 1c). In each PGG, the income of individual x will depend on the size of the group $k_x + 1$ (defined by the size of the neighbourhood centred on individual x ; see Fig. 1), on the number n_C of Cs in his/her neighbourhood and on the multiplication factor r applied to the group investment. The incomes of a defector and a cooperator in one group are given by $P_D = crn_C/(k_x + 1)$ and $P_C = P_D - c$, respectively, in the case where all cooperators contribute the same cost c per game, such that the contribution of an individual is proportional to his/her number of social ties. In the opposite limit, we considered the case in which C individuals with k_x neighbours contribute a cost $c/(k_x + 1)$ per game, such that the individual contribution of each C equals c independently of the number of social ties. Hence, the payoff of an individual y with a strategy s_y (1 if C, 0 if D) associated with the PGGs centred in a individual x is given by

$$P_{y,x} = \frac{r}{k_x + 1} \sum_{i=0}^{k_x} \frac{c}{k_i + 1} s_i - \frac{c}{k_y + 1} s_y$$

where $i = 0$ stands for x , s_i is the strategy of the neighbour i of x , and k_i is his/her degree.

Evolution. The network structure of the population defines not only the game interactions but also the structure through which strategy evolution proceeds. For non-repeated two-player games, this has been shown to constitute the most favourable model for cooperation³⁰. After engaging in all games, the accumulated payoff is mapped onto individual fitness. After each game round, all strategies are updated synchronously by following the finite population analogue of the replicator dynamics^{16,29}. When a site x with a payoff P_x is selected for update, a neighbour y (with a payoff P_y) is drawn at random between all k_x neighbours. If $P_x > P_y$, no update occurs. If $P_x < P_y$, x will adopt y 's strategy with a probability given by $(P_y - P_x)/M$. M ensures the proper normalization and is given by the maximum possible difference between the payoffs of x and y . The results are robust with respect to changes both in the detailed form of the normalization factor or if we adopt an asynchronous update instead of the synchronous one.

Simulations. The results were obtained for communities with $N = 10^3$ individuals and an average connectivity of $z = 4$. Each equilibrium fraction of cooperators was obtained by averaging more than 2,000 generations after a transient period of 10^5 generations. We started with 50% of Cs randomly placed on the graph. Each data point depicted in Fig. 2 corresponds to an average over 1,000 simulations; that is, 100 runs for 10 different realizations of the same class of graph. Finally, the distributions depicted in Fig. 4 were obtained by averaging the fitness distributions over 50 scale-free graphs with $z = 4$ and populations of 10^3 individuals, all cooperators.

30. Ohtsuki, H., Nowak, M. A. & Pacheco, J. M. Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs. *Phys. Rev. Lett.* **98**, 108106 (2007).

An internal thermal sensor controlling temperature preference in *Drosophila*

Fumika N. Hamada¹, Mark Rosenzweig¹, Kyeongjin Kang¹, Stefan R. Pulver¹, Alfredo Ghezzi¹, Timothy J. Jegla² & Paul A. Garrity¹

Animals from flies to humans are able to distinguish subtle gradations in temperature and show strong temperature preferences^{1–4}. Animals move to environments of optimal temperature and some manipulate the temperature of their surroundings, as humans do using clothing and shelter. Despite the ubiquitous influence of environmental temperature on animal behaviour, the neural circuits and strategies through which animals select a preferred temperature remain largely unknown. Here we identify a small set of warmth-activated anterior cell (AC) neurons located in the *Drosophila* brain, the function of which is critical for preferred temperature selection. AC neuron activation occurs just above the fly's preferred temperature and depends on dTrpA1, an ion channel that functions as a molecular sensor of warmth. Flies that selectively express dTrpA1 in the AC neurons select normal temperatures, whereas flies in which dTrpA1 function is reduced or eliminated choose warmer temperatures. This internal warmth-sensing pathway promotes avoidance of slightly elevated temperatures and acts together with a distinct pathway for cold avoidance to set the fly's preferred temperature. Thus, flies select a preferred temperature by using a thermal sensing pathway tuned to trigger avoidance of temperatures that deviate even slightly from the preferred temperature. This provides a potentially general strategy for robustly selecting a narrow temperature range optimal for survival.

Although the physiology of all cells is affected by temperature, the expression of temperature-activated members of the transient receptor potential (TRP) family (thermoTRPs) can make cell excitability highly temperature-responsive⁵. ThermoTRPs are cation channels with highly temperature-dependent conductances that participate in thermosensation from insects to humans⁵. The *Drosophila melanogaster* TRP channel dTrpA1 promotes larval heat avoidance⁶ and can be activated by warming in oocytes⁷. We addressed whether dTrpA1 contributes to the selection of a preferred temperature in the adult fly. When allowed to distribute along a thermal gradient for 30 min, wild-type *D. melanogaster* adults prefer ~25 °C (ref. 3), their optimal growth temperature⁸. Compared to wild-type controls, dTrpA1 loss-of-function mutant animals showed increased accumulation in the warmest (28–32 °C) regions of the gradient ($P < 0.0001$), but not in the coolest (18–22 °C) regions ($P = 0.5$) (Fig. 1a, b, and Supplementary Fig. 1). A dTrpA1 genomic minigene rescued the phenotype (Fig. 1a, b). Animals heterozygous for dTrpA1 loss-of-function mutations also preferred slightly elevated temperatures (Supplementary Fig. 2a). Thus, dTrpA1 function is important for determining thermal preference and specifically contributes to avoidance of warm regions.

If dTrpA1 was involved in thermotransduction, it should regulate the warmth responsiveness of thermosensors. As the identity of the adult *Drosophila* thermosensors was unknown, we examined dTrpA1 protein expression (using anti-dTrpA1 antisera⁶). dTrpA1 expression was

detected in three sets of previously uncharacterized cells in the brain: lateral cell (LC), ventral cell (VC) and AC neurons (Fig. 2a). dTrpA1 was also detected in the proboscis, but ablation studies detected no contribution of the proboscis to warmth avoidance (Supplementary Fig. 3). To focus on the neurons that contribute to thermal preference we examined where the rescuing dTrpA1 minigene restored dTrpA1 expression, and found that dTrpA1 expression was restored specifically within AC neurons, but not LC or VC neurons (Fig. 2b). This suggested that dTrpA1 expression in AC neurons—two pairs of neurons at the brain's anterior (Fig. 2c, d)—sufficed to restore thermal preference and that AC neurons might act as thermosensors.

Temperature responsiveness of AC neurons was examined using the fluorescent calcium indicator G-CaMP⁹. When exposed to increasing temperature, AC neurons showed robust increases in G-CaMP fluorescence, reflecting warmth-responsive increases in intracellular calcium (Fig. 2h, i and k). Ten out of the 27 AC neurons

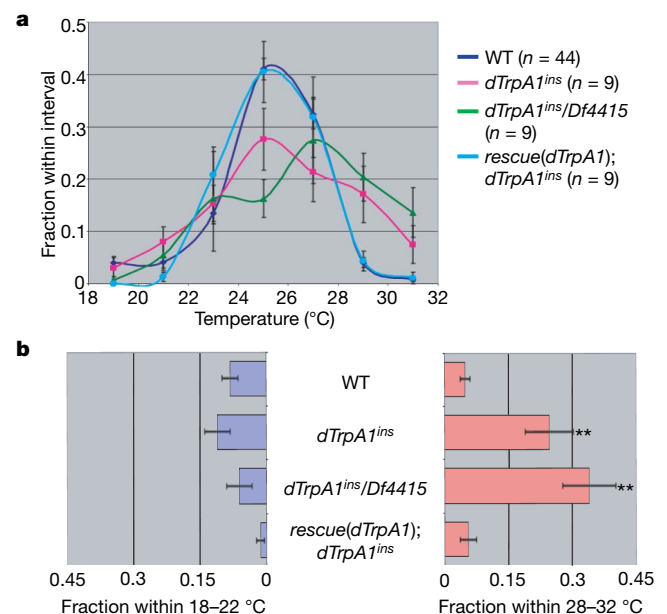


Figure 1 | dTrpA1 is required for warmth avoidance. **a**, Distribution of animals along a thermal gradient. WT, wild type. **b**, Fraction of animals in the 18–22 °C (blue) and 28–32 °C (red) regions of the thermal gradient. Data are mean \pm s.e.m; n represents the number of assays. Double asterisk, $P \leq 0.0001$ compared to wild type (unpaired t -test).

¹National Center for Behavioral Genomics, Volen Center for Complex Systems, Biology Department, Brandeis University MS-008, 415 South Street, Waltham, Massachusetts 02454, USA. ²Department of Cell Biology and Institute for Childhood and Neglected Diseases, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

imaged had fluorescence increases between 4% and 39%, with a mean increase over baseline ($\Delta F/F$) among these cells of 15% ($\pm 4\%$ s.e.m., $n = 10$; Fig. 2m). The average temperature at which fluorescence increases were initially observed was 24.9 °C (± 0.6 , $n = 10$), compatible with AC activation as temperatures rise above preferred. In contrast, none of the 21 *dTrpA1* mutant AC neurons imaged had fluorescence increases (Fig. 2j, l and m; $P < 0.003$ compared to wild

type, Fisher's exact test). As a control that mutant AC neurons remained physiologically active, we confirmed that they showed robust $\Delta F/F$ responses on potassium chloride addition (Supplementary Fig. 4). Notably, AC responses did not depend on an intact periphery, as all G-CaMP studies were performed using isolated brains from which peripheral tissues had been removed. These observations identify AC neurons as warmth-activated, *dTrpA1*-dependent thermosensors.

AC neurons project towards several brain regions, including the antennal lobe (Fig. 2d–g). The antennal lobe is implicated in cockroach thermosensation¹⁰, but has been studied exclusively for olfaction in *Drosophila*. So far, 11 of the ~50 antennal lobe glomeruli remain unassociated with identified olfactory receptors¹¹. AC neurites elaborated within two such unassociated glomeruli, VL2a and VL2p (Fig. 2f). Thus the *Drosophila* antennal lobe contains both thermosensory and olfactory neuron processes. VL2a is also innervated by Fruitless-expressing neurons implicated in pheromone transduction¹¹, suggesting that even individual glomeruli receive multi-modal sensory information. AC processes also branched within the suboesophageal ganglion and superior lateral protocerebrum, although these target regions are less defined than in the antennal lobe. These regions have been previously implicated in processing other types of sensory input¹².

As *dTrpA1* expression in AC neurons seemed sufficient to restore normal thermal preference, we examined whether such expression was also necessary. *dTrpA1* was knocked down selectively in AC neurons using tissue-specific RNA interference targeting *dTrpA1* controlled by *dTrpA1^{SH}-GAL4* (Fig. 3a), a promoter expressed in AC but not LC or VC neurons. Consistent with the importance of *dTrpA1* expression in AC neurons in thermal preference, AC knockdown increased the fraction of animals present in the 28–32 °C region compared to controls ($P \leq 0.0001$) (Fig. 3b, c). Similar results were obtained when *dTrpA1* expression was knocked down using a broad neuronal promoter (*Appl-GAL4*) (Fig. 3c and Supplementary Fig. 2b). All knockdowns were assessed by *dTrpA1* immunohistochemistry. *dTrpA1* knockdown with the general cholinergic neuron promoter *Cha(7.4)-GAL4* eliminated detectable *dTrpA1* expression in AC (and LC and VC) neurons, decreasing warmth avoidance (Fig. 3c and Supplementary Fig. 2c). In contrast, *dTrpA1* RNAi expressed using *Cha(1.2)-GAL4*—which is expressed in many brain cholinergic neurons¹³ but not AC neurons—did not disrupt warmth avoidance (Fig. 3c and Supplementary Fig. 2d). Taken together, our data suggest that *dTrpA1* expression in AC (but not LC or VC) neurons is both necessary and sufficient for normal thermal preference behaviour. Whether LC and VC neurons participate in other warmth-activated responses is unknown.

The identification of an internal sensor controlling temperature preference conflicts with the established view that *Drosophila* sense moderate warming using thermosensors in the third antennal segment³. We re-examined the effects of surgically removing either one third antennal segment and arista (unilateral ablation) or both (bilateral ablation). Both unilateral and bilateral ablation (see Methods) increased the fraction of animals in cool (18–22 °C), but not warm (28–32 °C), regions (Fig. 3d, f). Thus these tissues were dispensable for warmth avoidance, but essential for cool avoidance. When *dTrpA1* mutants were subjected to bilateral ablation, they accumulated in both cool and warm regions (Fig. 3e): the fraction between 18–22 °C did not differ from wild-type ablation animals ($P = 1.0$); the fraction between 28–32 °C did not differ from non-ablated *dTrpA1* mutants ($P = 0.9$) (Fig. 3f). Thus *dTrpA1*-expressing cells and antennal cells function additively to set preferred temperature, promoting avoidance of elevated and reduced temperatures, respectively.

These data are consistent with warmth activation of *dTrpA1* serving as the molecular basis of warmth sensing by AC neurons. As thermal activation of mammalian TRPA1 proteins is controversial, we tested whether *dTrpA1* could act as a molecular sensor of warming in the fly. Indeed, misexpression of *dTrpA1* throughout the fly nervous system (using *C155-GAL4*) caused a dramatic phenotype not

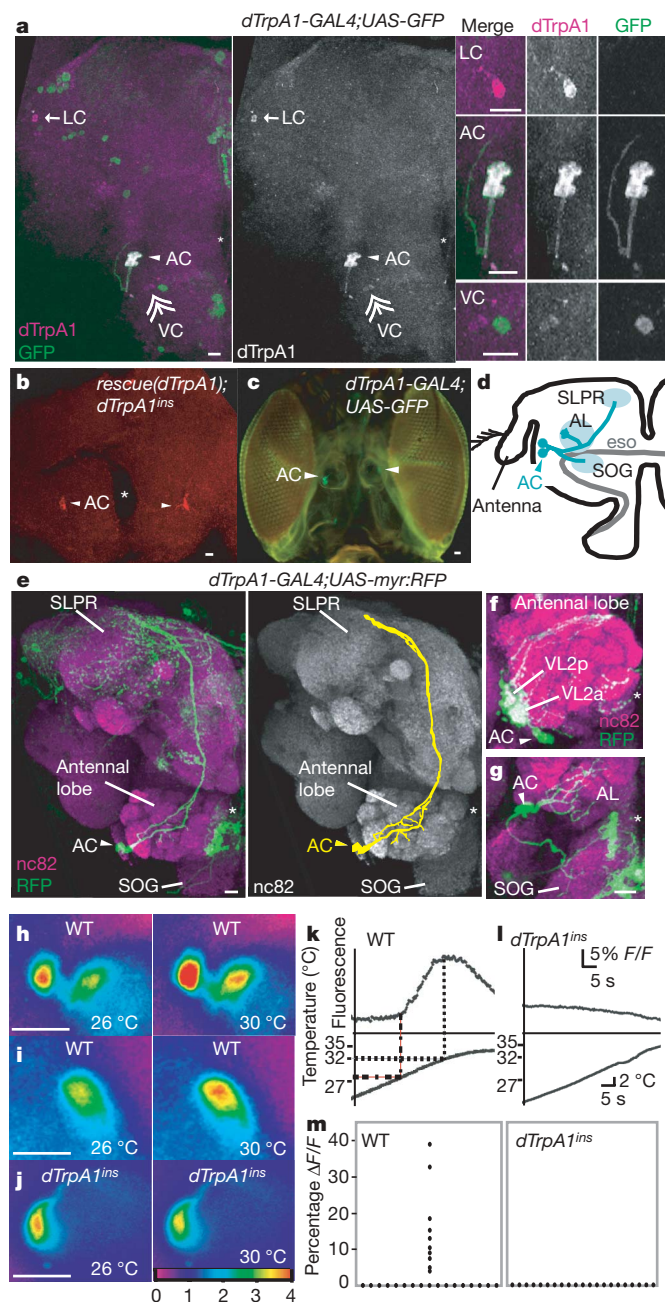


Figure 2 | AC neurons are thermosensors. **a**, AC (arrowhead), LC (arrow) and VC (double arrow) neurons are shown. **b**, *dTrpA1* minigene expression. **c**, AC neuron locations (antennae removed). **d**, AC neuron projections. **e**, Left, AC neuron processes labelled using *dTrpA1^{SH}-GAL4;UAS-myr::RFP*. Also labelled are cells that express *dTrpA1^{SH}-GAL4* but do not detectably express *dTrpA1* protein (also see **a**). Right, camera-lucida-style outline of AC neuron projections. **f**, **g**, AC neuron projections to the antennal lobe (AL; **f**) and suboesophageal ganglion (SOG; **g**). **h–j**, G-CaMP-labelled AC neurons. Two AC neurons are sometimes imaged simultaneously (**h**). Colour scale reflects fluorescence intensity (see Methods). **k**, **l**, Warmth-responsive G-CaMP fluorescence of AC neurons. **m**, Maximum $\Delta F/F$ of each AC neuron imaged. eso (also asterisk), oesophagus; GFP, green fluorescent protein; SLPR, superior lateral protocerebrum. Scale bars, 10 micrometres.

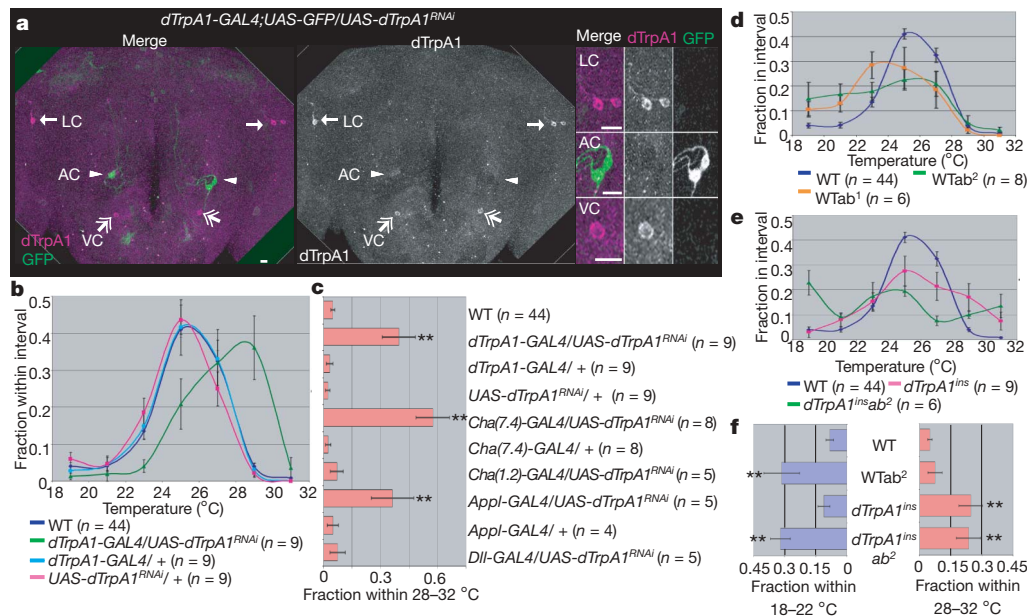
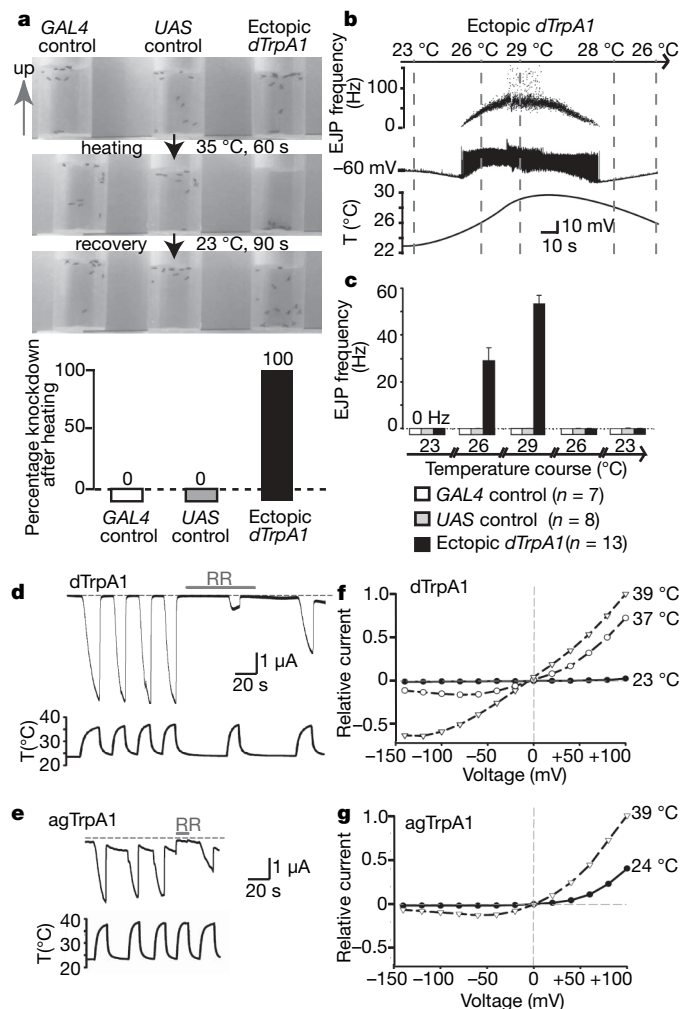


Figure 3 | AC neurons are necessary for warmth avoidance. **a**, AC-specific knockdown of dTrpA1 protein expression in *dTrpA1^{SH}-GAL4;UAS-GFP;UAS-dTrpA1^{RNAi}* animals. dTrpA1 is expressed in LC (arrow) and VC (double arrow) but not AC neurons (arrowhead). GFP marks dsRNA-expressing cells. The two left panels show adult brain and the right-hand panels show higher magnification of specific cells. **b**, Distribution of

indicated genotypes along a thermal gradient. **c**, Fraction of RNAi animals in 28–32 °C region of gradient. **d**, **e**, Animal distributions along a thermal gradient. **ab¹**, unilateral ablation; **ab²**, bilateral ablation. **f**, Fraction of animals in 18–22 °C (blue) and 28–32 °C (red) regions of gradient. Double asterisk, $P < 0.0001$. Error bars, s.e.m.; scale bars, 10 micrometres.



observed in controls: heating these flies to 35 °C for 60 s caused incapacitation, an effect reversed on return to 23 °C (Fig. 4a and Supplementary Movie 1). Similar effects were observed using electrophysiology, with moderate warming (above ~25 °C) triggering a barrage of excitatory junction potentials at the neuromuscular junction (Fig. 4b, c and Supplementary Fig. 5). These data strongly support dTrpA1 acting as a molecular sensor of warming. The ability of dTrpA1 mis-expression to confer warmth activation also suggests that dTrpA1 can be used as a genetically encoded tool for cell-specific, inducible neuronal activation. dTrpA1 might be particularly useful in tissues such as the fly brain where thermal stimulation is easier to deliver than the chemical or optical stimulation that controls other tools for modulating neuronal activity.

To test whether warmth activation is a property of other insect TrpA1s, we examined the malaria mosquito *Anopheles gambiae* TrpA1 (agTrpA1). As previously reported, dTrpA1 is warmth-activated when expressed in *Xenopus laevis* oocytes (Fig. 4d, f). We found that agTrpA1 also showed robust warmth activation (Fig. 4e, g). These currents were specific, they were not observed in uninjected oocytes (Supplementary Fig. 6) and were inhibited by ruthenium red (which antagonizes other TRPs). Similar to mammalian thermoTRPs, both dTrpA1 and agTrpA1 showed outward rectification (Fig. 4f, g). Closely related TrpA1s are present in the flour beetle *Tribolium castaneum* and in disease vectors such as *Pediculus humanus corporis*.

Figure 4 | dTrpA1 is a warmth sensor. **a**, Flies expressing dTrpA1 in all neurons (*C155-GAL4;UAS-dTrpA1*) are incapacitated after 60 s at 35 °C, but recover at 23 °C. Five experiments per genotype, 15 flies per experiment, s.e.m.s = 0. **b**, **c**, Warming (above ~25 °C) stimulates transmission at neuromuscular junction in *C155-GAL4;UAS-dTrpA1*. Ectopic dTrpA1, *C155-GAL4;UAS-dTrpA1*; EJP, excitatory junction potential; GAL4 control, *C155-GAL4; UAS control*, *UAS-dTrpA1*. Error bars, s.e.m. **d**–**g**, dTrpA1 and agTrpA1 are warmth-activated in oocytes (n > 14 each). **d**, **e**, Warmth-evoked currents in dTrpA1- or agTrpA1-expressing oocytes (–60 mV). Oocytes were injected with BAPTA 30 min before recording, minimizing cytosolic calcium elevations. RR, 50 μM ruthenium red; T, temperature. **f**, **g**, Current–voltage relationships of dTrpA1 and agTrpA1 at indicated temperatures.

(body lice), *Culex pipiens* (common house mosquito) and *Aedes aegypti* (yellow and dengue fever mosquito) which use warmth-sensing for host location and habitat selection^{14,15} (Supplementary Fig. 7). Such insect TrpA1s constitute potential targets for disrupting thermal preference and other thermosensory behaviours in agricultural pests and disease vectors.

Environmental temperature affects the physiology of all animals. Increasing temperatures associated with climate change are linked to poleward redistributions of hundreds of species including insects, fish, birds and mammals¹⁶. Although previously identified ambient thermoreceptors are peripheral^{4,17–20}, AC neurons are internal. As a ~1 mg fly is readily penetrated by ambient temperature variations^{21,22}, such an internal sensor should monitor environmental temperature effectively. dTrpA1 activation seems to be critical for AC neuron activation, suggesting that dTrpA1 threshold and expression changes could modulate thermal preferences. More speculatively, changes in insect TrpA1 function and expression could facilitate movements into novel environments or development of novel behaviours such as host seeking.

Although effects of environmental temperature on behaviour are ubiquitous, the mechanisms animals use to seek out optimal temperatures are largely unknown. AC neurons become active as temperatures rise above the preferred temperature, suggesting that they may function as 'discomfort' receptors that, together with putative antennal cool receptors (similar to those described in other insect antennae¹⁹), repel the fly from all but the most optimal temperatures. Notably, mice lacking the cool-activated channel TRPM8 prefer abnormally cool temperatures, whereas mice lacking heat-activated TRPV4 prefer warmer temperatures²³, indicating that similar strategies may be used in mammals.

METHODS SUMMARY

Fly strains and immunohistochemistry. The *dTrpA1* rescue transgene contains an NheI–XhoI genomic fragment (from 2.5 kb upstream to 1.2 kb downstream of the *dTrpA1* open reading frame) in pPelican containing Su(Hw) insulators²⁴. *dTrpA1^{SH}-GAL4* contains a GAL4 coding region flanked 5' by 2.5 kb upstream of the *dTrpA1* start codon and 3' by 1.7 kb of *dTrpA1* sequences from exon 1 to the fourth intron and inserted into pPelican. *dTrpA1^{SH}-GAL4* differs from the previously described *dTrpA1-GAL4* (ref. 6), and the two drivers overlap distinct subsets of dTrpA1-expressing cells. *dTrpA1^{ins}* was generated by ends-in gene targeting²⁵ and contains a tandem array of two mutated copies of *dTrpA1* as shown in Supplementary Fig. 1. One copy is deleted for the sixth transmembrane domain and carboxy terminus, whereas the other copy is deleted for the promoter region and start codon and contains a translational stop before the transmembrane domains. *Df(3L)ED4415* is a chromosomal deficiency that completely removes the *dTrpA1* locus. *dTrpA1* RNAi constructs were designed using the strategy of Kalidas and Smith²⁶ and generated as described²⁷. Further details are provided in Methods.

Temperature preference behaviour. The temperature preference behaviour assay was modified from that previously described³ as detailed in Methods.

Physiology. For G-CaMP imaging, brains from live *dTrpA1^{SH}-GAL4;UAS-G-CaMP* (wild type) or *dTrpA1^{SH}-GAL4;UAS-G-CaMP;dTrpA1^{ins}* flies were dissected in modified standard solution F (5 mM Na-HEPES, 115 mM NaCl, 5 mM KCl, 6 mM CaCl₂, 1 mM MgCl₂, 4 mM NaHCO₃, 5 mM trehalose, 10 mM glucose, 65 mM sucrose, pH 7.5)²⁸ by removing the cuticle from the head and trachea using fine forceps. The brain was severed from the body and placed onto sylgard-coated coverslips, anterior side facing up. Brains were imaged and oocyte physiology and larval recordings were performed as described in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 February; accepted 15 April 2008.

Published online 11 June 2008.

1. Fraenkel, G. & Gunn, D. *The Orientation of Animals. Kineses, Taxes and Compass Relations* (Clarendon, Oxford, 1940).
2. Fanger, P. O., Ostberg, O., Nicholl, A., Breum, N. O. & Jerking, E. Thermal comfort conditions during day and night. *Eur. J. Appl. Physiol.* **33**, 255–263 (1974).
3. Sayeed, O. & Benzer, S. Behavioral genetics of thermosensation and hygrosensation in *Drosophila*. *Proc. Natl Acad. Sci. USA* **93**, 6079–6084 (1996).

4. Mori, I. Genetics of chemotaxis and thermotaxis in the nematode *Caenorhabditis elegans*. *Annu. Rev. Genet.* **33**, 399–422 (1999).
5. Dhaka, A., Viswanath, V. & Patapoutian, A. TRP ion channels and temperature sensation. *Annu. Rev. Neurosci.* **29**, 135–161 (2006).
6. Rosenzweig, M. et al. The *Drosophila* ortholog of vertebrate TRPA1 regulates thermotaxis. *Genes Dev.* **19**, 419–424 (2005).
7. Viswanath, V. et al. Opposite thermosensor in fruitfly and mouse. *Nature* **423**, 822–823 (2003).
8. Siddiqui, W. H. & Barlow, C. A. Population growth of *Drosophila melanogaster* (Diptera: Drosophilidae) at constant and alternating temperatures. *Ann. Entomol. Soc. Am.* **65**, 993–1001 (1972).
9. Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise Ca²⁺ probe composed of a single green fluorescent protein. *Nature Biotechnol.* **19**, 137–141 (2001).
10. Fischer, H. & Tichy, H. Cold-receptor cells supply both cold- and warm-responsive projection neurons in the antennal lobe of the cockroach. *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.* **188**, 643–648 (2002).
11. Vosshall, L. B. & Stocker, R. F. Molecular architecture of smell and taste in *Drosophila*. *Annu. Rev. Neurosci.* **30**, 505–533 (2007).
12. Ito, K. et al. The organization of extrinsic neurons and their implications in the functional roles of the mushroom bodies in *Drosophila melanogaster* Meigen. *Learn. Mem.* **5**, 52–77 (1998).
13. Kitamoto, T., Ikeda, K. & Salvaterra, P. M. Regulation of choline acetyltransferase/lacZ fusion gene expression in putative cholinergic neurons of *Drosophila melanogaster*. *J. Neurobiol.* **28**, 70–81 (1995).
14. Brown, A. W. A. Factors in the attractiveness of bodies for mosquitoes. *Nature* **167**, 202 (1951).
15. Friend, W. G. & Smith, J. J. B. Factors affecting feeding by bloodsucking insects. *Annu. Rev. Entomol.* **22**, 309–331 (1977).
16. Parmesan, C. & Yohe, G. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**, 37–42 (2003).
17. Patapoutian, A., Peier, A. M., Story, G. M. & Viswanath, V. ThermoTRP channels and beyond: mechanisms of temperature sensation. *Nature Rev. Neurosci.* **4**, 529–539 (2003).
18. Tominaga, M. & Caterina, M. J. Thermosensation and pain. *J. Neurobiol.* **61**, 3–12 (2004).
19. Tichy, H. & Gingl, E. Problems in hygro- and thermoreception. In *The Ecology of Sensing* (eds Barth, F. G. & Schmid, A.) 271–287 (Springer, New York, 2001).
20. Liu, L., Yermolaieva, O., Johnson, W. A., Abboud, F. M. & Welsh, M. J. Identification and function of thermosensory neurons in *Drosophila* larvae. *Nature Neurosci.* **6**, 267–273 (2003).
21. Stevenson, R. D. Body size and limits to the daily range of body temperature in terrestrial ectotherms. *Am. Nat.* **125**, 102–117 (1985).
22. Heinrich, B. *The Hot-Blooded Insects: Strategies and Mechanisms of Thermoregulation* (Harvard Univ. Press, Cambridge, USA, 1993).
23. Caterina, M. J. Transient receptor potential ion channels as participants in thermosensation and thermoregulation. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **292**, R64–R76 (2007).
24. Barolo, S., Castro, B. & Posakony, J. W. New *Drosophila* transgenic reporters: insulated P-element vectors expressing fast-maturing RFP. *Biotechniques* **36**, 436–440, 442 (2004).
25. Rong, Y. S. & Golic, K. G. Gene targeting by homologous recombination in *Drosophila*. *Science* **288**, 2013–2018 (2000).
26. Kalidas, S. & Smith, D. P. Novel genomic cDNA hybrids produce effective RNA interference in adult *Drosophila*. *Neuron* **33**, 177–184 (2002).
27. Tayler, T. D., Robichaux, M. B. & Garrity, P. A. Compartmentalization of visual centers in the *Drosophila* brain requires Slit and Robo proteins. *Development* **131**, 5935–5945 (2004).
28. Ng, M. et al. Transmission of olfactory information between three populations of neurons in the antennal lobe of the fly. *Neuron* **36**, 463–474 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank Garrity laboratory members and L. Griffith, L. Huang, R. Huey, E. Marder, C. Miller, M. Rosbash, P. Sengupta, G. Turrigiano and their laboratories for advice and manuscript comments. Supported by NINDS (PO1 NS044232, P30 NS045713S10 and RR16780), NEI (RO1 EY13874, P.A.G.), NIMH (RO1 MH067284, to L. Griffith (for S.P. and A.G.)), and Japan Society for the Promotion of Science (F.N.H.).

Author Contributions F.N.H., M.R., S.R.P., K.K. and P.A.G. designed experiments; F.N.H. performed behaviour and imaging; M.R. created the *dTrpA1* mutant, *Gal4*, RNAi and rescue strains; K.K. performed oocyte electrophysiology and dTrpA1 overexpression; S.R.P. performed NMJ electrophysiology, A.G. assisted with imaging; T.J.J. isolated agTrpA1 cDNA; and P.A.G. performed bioinformatics and assisted with knockdown studies. F.N.H. and P.A.G. wrote the paper with assistance from M.R., K.K., S.R.P. and A.G.

Author Information The sequence for agTrpA1 has been deposited in the GenBank database under accession number EU624401. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.A.G. (pgarrity@brandeis.edu).

METHODS

Fly strains. The *dTrpA1* rescue transgene contained an NheI–XhoI genomic sequence fragment (extending from 2.5 kb upstream of the *dTrpA1* start codon to 1.2 kb downstream of the stop codon) and was inserted into a pPelican vector containing Su(Hw) insulator sequences²⁴. *dTrpA1^{SH}-GAL4* contains a Gal4 coding region flanked 5' by 2.5 kb upstream of the *dTrpA1* start codon and 3' by 1.7 kb of *dTrpA1* sequences from exon 1 to the fourth intron and inserted into pPelican. *dTrpA1^{SH}-GAL4* differs from the previously described *dTrpA1-GAL4* (ref. 6) as it contains more *dTrpA1* sequences as well as flanking Su(Hw) insulator sequences, and the two drivers overlap distinct subsets of *dTrpA1*-expressing cells. The *dTrpA1^{ins}* mutant was generated by ends-in homologous recombination-mediated gene targeting²⁵ and contains a tandem array of two mutated copies of *dTrpA1*: one copy lacks DNA encoding the predicted sixth transmembrane domain and the C terminus, whereas the other lacks DNA sequences containing the putative promoter region and predicted start codon and carries an insertion creating a translational stop before the transmembrane domains. The targeting construct was generated by cloning a BglI–KpnI *dTrpA1* genomic DNA fragment from BAC RP98-10P9 (Open Biosystems) into pBluescriptII. An I-SceI site was inserted into the BamHI site in exon 8 of *dTrpA1*, a ClaI site in exon 4 was filled in to generate the engineered frameshift mutation, and the construct was cloned into NotI/KpnI digested pTV2 for targeting as described²⁵. The organization of the *dTrpA1^{ins}* mutant locus is described in Supplementary Fig. 1. *Df(3L)4415* is a chromosomal deficiency that completely removes the *dTrpA1* locus. *dTrpA1* RNAi constructs were designed using the strategy of Kalidas and Smith²⁶ and generated as described²⁷. *UAS-dTrpA1^{RNAi}* flies contain two copies each of *UAS-dTrpA1^{RNAi}-A* and *UAS-dTrpA1^{RNAi}-B*. The PCR primers used to create *UAS-dTrpA1^{RNAi}-A* were genomic fragment 5'-ATAACTGAGTTCGATGCATGCCACG-3' and 5'-GCCTCGAGACTAGTCTGGAAAATGGAAGCCAAAGT-3'; cDNA fragment 5'-GCTCTAGAATAACTGAGTTCGATGCATGCCACG-3' and 5'-CACTC GAGACTAGTCTGTTTCCAACCGCTACGAG-3'. The PCR primers used to create *UAS-dTrpA1^{RNAi}-B* were genomic fragment 5'-AGGAGCGGGC CAACGAGGTGATG-3' and 5'-GCCTCGAGACTAGTCTGAAAATGGAGG TGTGCTATATG-3'; cDNA fragment 5'-ATTCTAGAAGGAGCGGGCCAA CGAGGTGATG-3' and 5'-GCCTCGAGACTAGTCCCATATTGAGTATTGA CTCATC-3'.

Further fly strains were obtained from Bloomington, except for *Cha(7.4)-GAL4* and *Cha(1.2)-GAL4* which were from P. Salvaterra, *Appl-GAL4* was from T. Tayler, *Dll-GAL4* was from G. Boekx and *UAS-G-CaMP* was from R. Axel.

Temperature preference behaviour. A temperature preference behaviour assay was designed as previously described³ but with a larger temperature range. Air temperatures inside the apparatus were determined using a Fluke 52II thermometer with multiple temperature probes. The apparatus was coated with Rain-X to prevent flies from escaping the temperature gradient; this does not disturb the normal temperature preference (data not shown). For each assay, ~20–30 adult flies (0–3 days old) were blown into the apparatus through a hole at the midpoint of the gradient and exposed to the gradient for 30 min in darkness before data collection. All experiments were performed in an environmental room maintained at 25 °C, 70% relative humidity.

We plotted performance as a function of the air temperature within the behavioural chamber. We monitored both air and apparatus surface temperatures but found that relying on air temperature significantly decreased variation in behaviour between experiments, suggesting that this parameter was more relevant to the fly (F.N.H. and P.A.G., unpublished observations). As the surface temperature gradient is steeper than the air temperature gradient (F.N.H. and P.A.G., unpublished observations), our thermal gradients extended to both warmer and cooler temperatures than the gradients previously described³ (~31.5 °C in surface temperature is ~29 °C in air temperature). Our finding that adults lacking third antennal segments avoid warm regions of the thermal gradient was initially unexpected, as removal of the third antennal segment was previously reported to eliminate temperature preference³. However, these authors did note that the distribution of ablated flies fell off significantly at higher temperatures, approaching zero near 31.5 °C surface temperature. The apparent differences in our results may reflect, in part, subtle differences in assay environment including the monitoring of only surface temperature in the previous study³. Irrespective of subtle differences in our data sets, our ablation data together with our *dTrpA1* mutant analysis clearly demonstrate that avoidance of warm temperatures persists in the absence of the third antennal segment. Our conclusion that unilateral ablation causes a partial decrease in cold avoidance also differs from the conclusion of the previous report that unilateral ablation has no effect³; however, examination of the data in the earlier study³ also indicates that a significant (>3-fold) increase in the fraction of flies in the 18–22.5 °C region (from ~15% to ~50%) occurred following unilateral ablation.

Immunostaining. Immunostaining was performed as described⁶ except 5% normal goat serum and 1% bovine serum albumin in PBST (PBS plus 1% Triton X-100) were used for blocking and antibody incubations, and PBST was used for washing. The antibodies used were: rat anti-dTrpA1 (ref. 6; 1:1,000), rabbit anti-RFP (1:200; Chemicon), mouse nc82 (1:40; Developmental Studies Hybridoma Bank), goat anti-rat Cy3 (1:4,000; Jackson ImmunoResearch), goat anti-rat Cy5 (1:200; Jackson ImmunoResearch), goat anti-rabbit Cy3 (1:200; Jackson ImmunoResearch) and goat anti-mouse Cy5 (1:200; Jackson ImmunoResearch).

Calcium imaging; fly preparation and imaging. Brains from live *dTrpA1^{SH}-GAL4;UAS-G-CaMP* (wild type) and *dTrpA1^{SH}-GAL4;UAS-G-CaMP;dTrpA1^{ins}* flies were dissected in modified standard solution F (5 mM Na-HEPES, 115 mM NaCl, 5 mM KCl, 6 mM CaCl₂, 1 mM MgCl₂, 4 mM NaHCO₃, 5 mM trehalose, 10 mM glucose, 65 mM sucrose, pH 7.5)²⁸ by removing the cuticle from the head and trachea using fine forceps. The brain was severed from the body and placed onto Sylgard-coated coverslips, anterior side facing up. The brain was then glued to the slide by applying small amounts (3 M) of Vetbond between the external edge of the optic lobes and the slide using a pulled glass capillary. The slide was mounted on a laminar flow perfusion chamber (~500 µl volume) beneath a ×40 or ×60 water immersion objective of a fixed-stage upright microscope (Olympus BX51W1), illuminated using a 75 W xenon Apo lamp with a 490 nm excitation filter and visualized through a 528 nm emission filter (Olympus). During experiments, the preparation was constantly perfused with modified standard solution F via gravity flow at a rate of ~3 ml min⁻¹. The solution temperature was gradually increased from room temperature (~22 °C) to 33 °C using a CL-100 bipolar temperature controller equipped with an SC-20 dual in-line solution heater/cooler (Warner Instruments). Optical images of the preparation were acquired during the temperature shift using a digital CCD camera (C4742-80-12AG; Hamamatsu) at 4 frames per second with 512 × 512 pixel resolution. The image data was digitized and analysed using Volocity software (Improvision). For analysis, areas representing G-CaMP-expressing cell bodies were circumscribed and the mean fluorescent intensity was calculated for each region of interest at every frame. Background fluorescence (calculated from the average fluorescence of two randomly chosen non-G-CaMP expressing areas) was subtracted from the mean fluorescent intensity of the regions of interest. Background-subtracted values were then expressed as percentage ΔF/F, where F is the mean fluorescence intensity in the 10 s before stimulation. The solution temperature was simultaneously recorded and digitized using PowerLab 4/30 and Chart software (AD Instruments) and synchronized with the image acquisition through an orbit II controller (Improvision). Bleaching corrections were done by plotting a least-squares fit line in Excel using the first 10 s of imaging and extrapolating this bleaching rate for the duration of the experiment. Pseudo-colour images were generated in Adobe Photoshop from pixel greyscale values by setting black level values to 40 and converting to pseudo-colour using the Spectrum tool.

Oocyte electrophysiology. Oocyte-positive *Xenopus laevis* females were obtained from Nasco, and maintained at ~19 °C with 12 h/12 h dark/light cycles. Ovaries were surgically isolated, and treated with 1.5 mg ml⁻¹ of collagenase type II (Worthington) for 90 min, and individual oocytes were defolliculated using forceps. Capped mRNA was produced using mMMessage mMachine T3 kit (Ambion) from each complementary DNA construct, and a 50-µl aliquot per oocyte was injected using an automatic Drummond microinjector. Membrane potential was maintained at -60 mV by two electrode voltage clamping during warm-activated current recording. Resistance of pulled glass capillary electrodes filled with 3 M KCl was between 0.5 and 1.5 MΩ. The typical resting membrane potentials of oocytes used for current recording were between -25 and -60 mV. Current was recorded at 2 kHz and filtered to 1 kHz with the output filter of the amplifier (OC-725B, Warner instruments). The temperature of the oocyte perfusion buffer (96 mM NaCl, 1 mM MgCl₂, 4 mM KCl, 5 mM HEPES, pH 7.6) was changed by SC-20 in-line heater/cooler (Warner Instruments) under the control of CL-100 bipolar temperature controller (Warner Instruments). Where indicated, 50 µl of 20 mM BAPTA was injected into oocytes 30 min before recording to minimize elevations in cytosolic calcium concentration. pClamp 8.0 and Sigmplot 8.0 were used to acquire and analyse the data. Voltage steps (-140 mV to 100 mV) were applied to assess current-voltage relationship of dTrpA1- or AgTrpA1-expressing oocytes. Each voltage step lasted 100 ms after 10 ms holding at -60 mV, and the recorded current in the last 40 ms was averaged to determine the current amplitude at a given voltage.

Larval electrophysiology. All larval dissections and physiological recordings were performed in HL3.1 physiological saline containing (in mM) 70 NaCl, 5 KCl, 0.8 CaCl₂, 4 MgCl₂, 10 NaHCO₃, 5 trehalose, 115 sucrose, 5 HEPES, pH 7.1–7.2; the osmolarity was ~309 mmol kg⁻¹.

Female third instar larvae were dissected and pinned dorsal side up in a Sylgard-lined Petri dish (Dow Corning). The gut and trachea were removed with

fine forceps, exposing the larval body wall muscles. The anterior lobes of the larval brain were removed, but the rest of the ventral ganglion was left undisturbed, preserving the cell bodies of motor neurons innervating body wall muscles.

Larval preparations were mounted on the stage of a BX50WI compound microscope (Olympus) and continuously superfused with HL3.1 using a custom built gravity-fed perfusion system. Temperature ramps were performed by circulating hot and then cold water past coils of perfusion tubing sealed in a PVC pipe on the way to the preparation. Bath temperature was ramped from 23 °C to 29 °C then back to 23 °C in every experiment; ramp time was typically 3–5 min each way. Bath temperatures were monitored with a Physitemp model BAT-12 thermometer with thermocouple probe or alternatively using an SC-20 in-line heater/cooler (Warner Instruments) under the control of CL-100 bipolar temperature controller (Warner Instruments).

Larval muscle 6 was targeted for intracellular work. Excitatory junction potential frequency in muscle 6 was monitored in control and experimental animals as bath temperatures rose and fell. Recordings from muscle 6 were performed with sharp glass electrodes (12–18 M Ω) filled with 3 M KCl. Voltage signals were amplified with either an Axoclamp 2A or Axopatch 200B (Axon Instruments) and digitized with a Powerlab 4/30 data acquisition system (ADInstruments). Voltage traces were recorded in Chart 5.1 (ADInstruments). The data were analysed using scripts in Spike 2 (version 5, CED) and standard features in Microsoft Excel.

Osteoclast size is controlled by Fra-2 through LIF/LIF-receptor signalling and hypoxia

Aline Bozec^{1†}, Latifa Bakiri^{1†}, Astrid Hoebertz¹, Robert Eferl^{1†}, Arndt F. Schilling², Vukoslav Komnenovic¹, Harald Scheuch¹, Matthias Priemel², Colin L. Stewart³, Michael Amling² & Erwin F. Wagner^{1†}

Osteoclasts are multinucleated haematopoietic cells that resorb bone. Increased osteoclast activity causes osteoporosis, a disorder resulting in a low bone mass and a high risk of fractures¹. Increased osteoclast size and numbers are also a hallmark of other disorders, such as Paget's disease and multiple myeloma². The protein c-Fos, a component of the AP-1 transcription factor complex, is essential for osteoclast differentiation³. Here we show that the Fos-related protein Fra-2 controls osteoclast survival and size. The bones of Fra-2-deficient newborn mice have giant osteoclasts, and signalling through leukaemia inhibitory factor (LIF) and its receptor is impaired. Similarly, newborn animals lacking LIF have giant osteoclasts, and we show that LIF is a direct transcriptional target of Fra-2 and c-Jun. Moreover, bones deficient in Fra-2 and LIF are hypoxic and express increased levels of hypoxia-induced factor 1 α (HIF1 α) and Bcl-2. Overexpression of Bcl-2 is sufficient to induce giant osteoclasts *in vivo*, whereas Fra-2 and LIF affect HIF1 α through transcriptional modulation of the HIF prolyl hydroxylase PHD2. This pathway is operative in the placenta, because specific inactivation of Fra-2 in the embryo alone does not cause hypoxia or the giant osteoclast phenotype. Thus placenta-induced hypoxia during embryogenesis leads to the formation of giant osteoclasts in young pups. These findings offer potential targets for the treatment of syndromes associated with increased osteoclastogenesis.

Little is known about the role of the Fos-related protein Fra-2 (encoded by *Fosl2*) in osteoclastogenesis, and *Fosl2*^{-/-} pups die within a week after birth^{4,5}. Detailed analysis revealed severe osteopenia as early as embryonic day (E)18.5 (data not shown). Von Kossa staining and quantitative histomorphometry of postnatal bones showed a 50% decrease in mineralized bone, whereas osteoblast numbers were not altered (Fig. 1a and Supplementary Fig. 1a, b). Most strikingly, staining for the osteoclast markers tartrate-resistant acid phosphatase (encoded by *TRAP*) and cathepsinK (encoded by *CathK*) showed a marked increase in both the number and size of Fra-2-deficient osteoclasts (Fig. 1a and Supplementary Fig. 1a). Quantification showed that the number, the relative surface covered by osteoclasts and the mean osteoclast surface were all increased—threefold, fivefold and twofold, respectively. Consistently, serum deoxypyridinoline crosslinks indicated increased resorptive activity in mutant newborn animals (Supplementary Fig. 1b), and quantitative reverse-transcriptase-mediated polymerase chain reaction (qPCR) revealed a significant increase in osteoclast marker gene expression, such as *TRAP*, *CathK*, *RANK*, *cfrms* and *MITF* (Supplementary Fig. 1c, d). Among the genes affecting osteoclast size and numbers, *LIFR* (encoding the LIF receptor)⁶ was found to be markedly decreased in Fra-2-deficient bones (Supplementary Fig. 1c).

Interleukin (IL)-6-type cytokines such as leukaemia inhibitory factor (LIF) and oncostatin (Osm) share the gp130 co-receptor and regulate osteoclastogenesis by means of the JAK/Stat or MAPK/ERK pathways⁷. *LIFR*, *gp130* and *Stat-1* knockouts have larger and more numerous osteoclasts^{6,8}. *LIF* and *Stat-1* mRNA were reduced in Fra-2-deficient bones, whereas *Osm* and *Stat-3* were increased and *IL-6*, *CNTF*, *gp130*, *IL-6R* and *OsmR* were unchanged (Supplementary Fig. 1e). Downregulation of *LIFR*, LIF and phosphorylated Stat-1 was confirmed at the protein level, whereas phosphorylated Stat-3 (P-Stat-3) was unchanged and phosphorylated ERK1/2 was increased (Fig. 1b and Supplementary Fig. 1f). p38, phosphorylated p38 and phosphorylated AKT were unaffected (Supplementary Fig. 1g). Immunohistochemistry indicated that *LIFR*, LIF and P-Stat-1 were decreased in all cell types of *Fosl2* knockout long bones (Supplementary Fig. 1h).

LIF mRNA expression is undetectable in primary osteoclasts, but was significantly decreased in *Fosl2*^{-/-} stromal cell cultures (data not shown). To analyse whether *LIF* is a transcriptional target of Fra-2, chromatin immunoprecipitation (ChIP) was performed with anti-Fra-2 antibodies and wild-type and *Fra-2*-deficient osteoblasts. Two DNA fragments, a proximal fragment within a previously described LIF enhancer⁹ and a distal fragment, were specifically amplified from *Fosl2* wild-type immunoprecipitates (Fig. 1c, d). The distal fragment was immunoprecipitated with antibodies against Jun proteins, regardless of whether Fra-2 was expressed or not (Fig. 1d). Binding to the proximal fragment was restricted to Fra-2 and c-Jun in wild-type cells, whereas no c-Jun binding was detected in the absence of Fra-2 (Fig. 1d). ChIP assays with antibodies against methylated histone H3 (ref. 10) confirmed that the distal fragment was not transcriptionally active in wild-type cells. However, a striking decrease in active (H3K4me3) methylation marks and a marked increase in repressive (H3K27me3) methylation marks was observed at the proximal fragment in *Fosl2* knockout immunoprecipitates (Supplementary Fig. 1i). Moreover, increased activity of a luciferase reporter containing the proximal fragment was measured, when using c-Jun and Fra-2 expression vectors, as well as with a combination of both (Fig. 1e). Finally, a c-Jun~Fra-2 forced dimer¹¹ efficiently transactivated the same reporter, whereas a reporter harbouring a mutated AP-1-binding site (mLIFP-Luc) was unaffected (Fig. 1e). These data indicate that Fra-2 and c-Jun form heterodimers to regulate *LIF* gene transcription positively.

LIF knockout mice are viable, and no bone phenotype has been reported^{12,13}. Von Kossa staining and histomorphometry revealed a 40% decrease in bone volume in *LIF*-mutant newborn animals, whereas osteoblasts were not altered (Fig. 2a and Supplementary Fig. 2a). Giant osteoclasts were identified by TRAP staining

¹Research Institute of Molecular Pathology (I.M.P.), Dr. Bohr-Gasse 7, A-1030 Vienna, Austria. ²Center for Biomechanics and Skeletal Biology, Department of Trauma, Hand, and Reconstructive Surgery, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, D-20246 Hamburg, Germany. ³Developmental and Regenerative Biology Laboratory, Institute of Medical Biology (I.M.B.), 8A Biomedical Grove, #06-06 Immunos, 138648 Singapore. †Present addresses: Spanish National Cancer Research Centre (C.N.I.O.), C/ Melchor Fernández Almagro, 3, E-28029 Madrid, Spain (A.B., L.B., E.F.W.); Ludwig Boltzmann Institute for Cancer Research, Währinger Strasse 13a, A-1090 Vienna, Austria (R.E.).

(Fig. 2a) and osteoclast numbers and size as well as relative osteoclast surface and bone resorption were significantly increased (Supplementary Fig. 2b). There was a consistent increase in *TRAP*, *CathK* and *cfms* mRNA, although *RANK* and *MITF* were unchanged (Supplementary Fig. 2c). Interestingly, decreased *Fosl2* expression was observed in LIF-deficient bones, suggesting a regulatory loop operating between LIF/LIFR signalling and Fra-2 (Supplementary Fig. 2c). *Osm* and *LIFR* expression was decreased, whereas expression of the other IL-6/gp130 genes was not affected (Supplementary Fig. 2d, e). Similarly to Fra-2-deficient bones, *Stat-1* expression was diminished, whereas that of *Stat-3* was increased (Supplementary Fig. 2d).

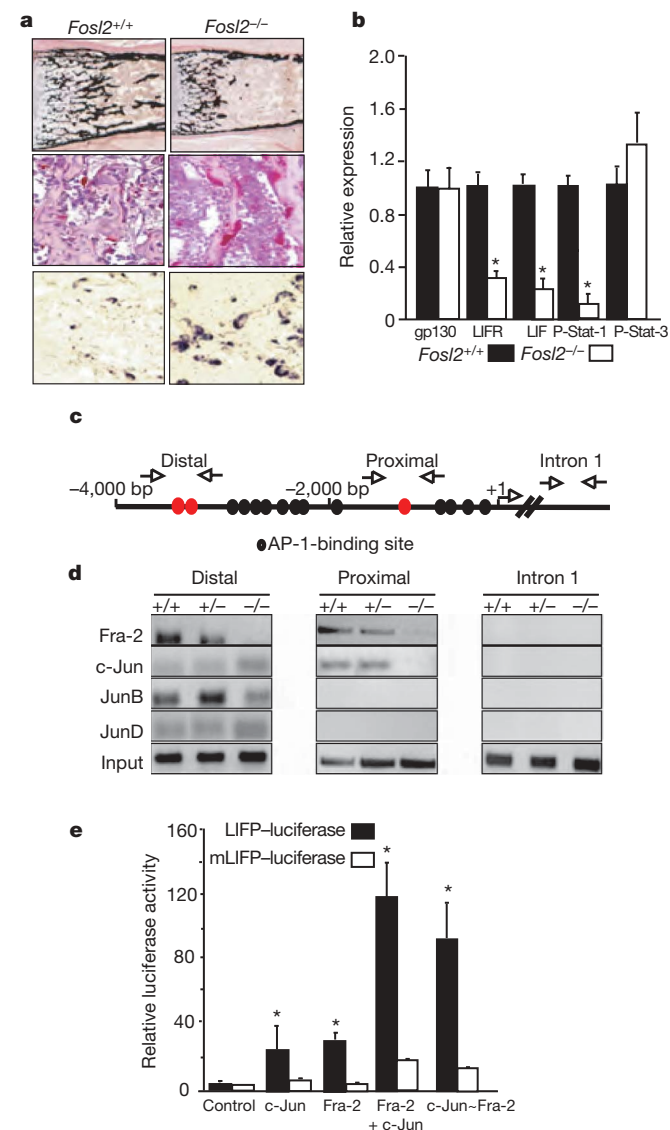


Figure 1 | Giant osteoclasts and decreased LIF/LIFR signalling in pups lacking Fra-2 (*Fosl2*). **a**, Von Kossa (top) and TRAP (middle) staining (at postnatal day (P)3) and *in situ* hybridization for *CathK* (bottom, P2) in *Fosl2*^{+/+} and *Fosl2*^{-/-} tibiae. Original magnifications $\times 5$ (top) and $\times 40$ (middle and bottom). **b**, Western blot analyses of gp130, LIFR, LIF, P-Stat-1 and P-Stat-3 in P2 long bones ($n = 5$, relative to actin, wild type set to 1). **c**, Putative (black) and Fra-2-binding (red) AP-1 sites in the *LIF* enhancer. Arrows indicate primers amplifying distal and proximal fragments and intron 1. bp, base pairs. **d**, Chromatin of the indicated genotypes was immunoprecipitated with AP-1 antibodies. Endpoint qPCR-amplified fragments are shown. **e**, Wild-type (LIFP-luciferase) or AP-1-mutated (mLIFP-luciferase) reporter assay for the proximal *LIF* fragment in the presence of c-Jun, Fra-2, Fra-2 + c-Jun or c-Jun~Fra-2 expression vectors ($n = 3$). Error bars represent s.d.; asterisk, $P < 0.01$.

Fetal liver-derived osteoclast progenitors deficient in Fra-2 or LIF failed to differentiate *in vitro*. Smaller and less numerous TRAP-positive cells were formed, with a reduced expression of osteoclast markers (Fig. 2b, c, Supplementary Fig. 2f, and data not shown). The addition of LIF completely rescued the differentiation defect and the number of osteoclasts containing more than five nuclei was identical to that in wild-type cultures (Fig. 2b, c). Because LIF is involved in the survival of myeloid progenitors *in vitro*¹², TdT-mediated dUTP nick end labelling (TUNEL) analyses revealed increased apoptosis of cells deficient in Fra-2 or LIF, which was efficiently rescued by the addition of LIF (Fig. 2b, c).

The osteoclast survival gene *Bcl-2* (refs 14, 15) was found significantly increased in bones deficient in Fra-2 or LIF, whereas the expression of other *Bcl-2*-related genes was unchanged

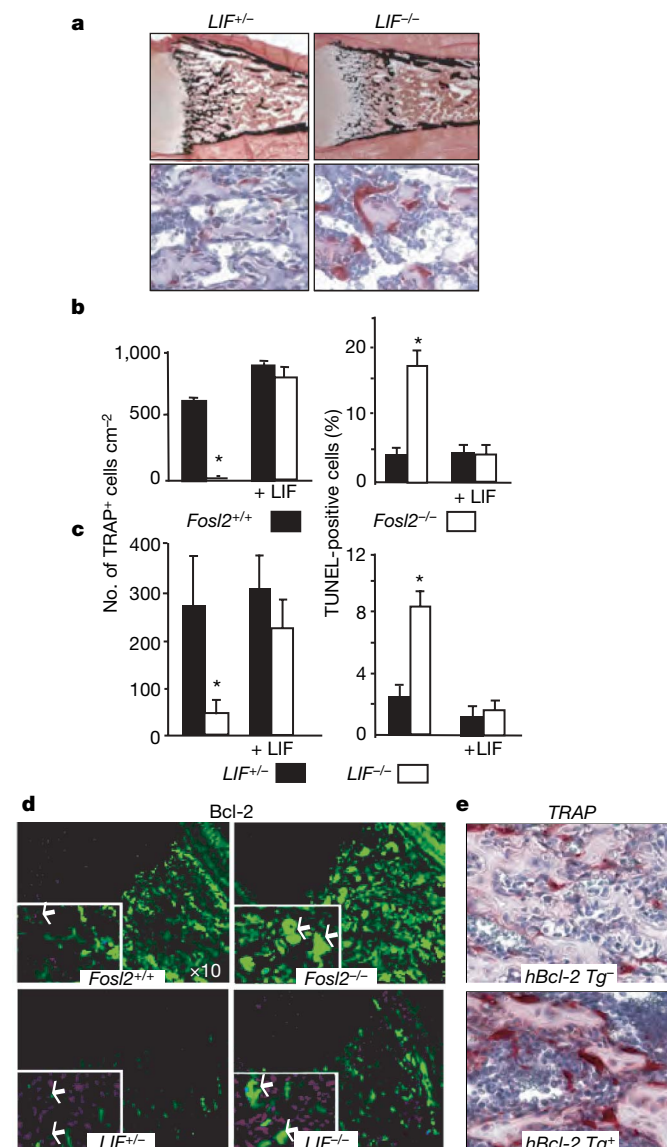


Figure 2 | LIF inactivation or Bcl-2 overexpression leads to giant osteoclasts. **a**, Von Kossa and TRAP staining of *LIF*^{+/+} and *LIF*^{-/-} tibiae at P2. Original magnifications $\times 5$ (top) and $\times 40$ (bottom). **b**, **c**, Impaired osteoclast differentiation *in vitro* (7 days, $n = 3$) with increased apoptosis (48 h, $n = 3$) and rescue by LIF of Fra-2-deficient (**b**) and LIF-deficient (**c**) fetal liver progenitors. Error bars represent s.d.; asterisk, $P < 0.01$. **d**, Bcl-2 immunohistochemistry in *Fosl2*^{+/+}, *Fosl2*^{-/-}, *LIF*^{+/+} and *LIF*^{-/-} P2 long bones. Original magnifications $\times 10$ and $\times 40$ (insets). Arrows in insets indicate osteoclasts. **e**, TRAP staining of wild-type (top) and *hBcl-2* Tg⁺ (bottom) tibiae at P2. Original magnification $\times 40$.

(Supplementary Fig. 3a, b). Immunohistochemistry showed that Bcl-2 was localized to the monocyte/osteoclast compartment (Fig. 2d and Supplementary Fig. 3c, d). Next, *Vav2*-human (*h*)*Bcl-2* transgenic mice overexpressing Bcl-2 in the haematopoietic lineage¹⁶ were analysed (Supplementary Fig. 3e). Decreased bone mass, increased osteoclast size and increased TRAP expression were detected (Fig. 2e and Supplementary Fig. 3f–h). Increased Bcl-2 expression is therefore sufficient to induce giant osteoclasts *in vivo*. Because progenitors deficient in Fra-2 or LIF were less viable *in vitro*, we proposed that signals from the bone microenvironment are responsible for Bcl-2 induction and extended osteoclast survival *in vivo*.

Hypoxia can induce Bcl-2 (ref. 17) and has also been described as a major stimulator of osteoclast formation and bone resorption^{18,19}. Hypoxic osteoclasts were detected in Fra-2 and LIF mutant long bones¹⁹ (Fig. 3a and Supplementary Fig. 4a). Immunohistochemistry revealed HIF1 α localization in both mutant osteoclasts (Fig. 3b), consistent with the hypothesis that environmental hypoxia can trigger the giant osteoclast phenotype. Increased *HIF1 α* RNA and HIF1 α protein were measured in both mutant bones (Supplementary Fig. 4b, c) and HIF1 α targets such as *bNIP3*, *iNOS* and *VEGF* were upregulated, whereas *Glut1* and *Glut3* were unchanged (Supplementary Fig. 4c, d). Increased expression of *HGF*, *c-Met* and *CD44*, which have been described to enhance osteoclast formation and survival²⁰ was also observed (Supplementary Fig. 4c). No changes in HIF1 α or HIF1 α targets were observed in the *Vav2*-*hBcl-2* transgenic bones, indicating that increased HIF1 α expression in osteoclasts is not due to increased osteoclastogenesis (Supplementary Fig. 3h).

To address the question of whether Fra-2 inactivation in osteoclast precursors is sufficient to trigger giant hypoxic osteoclasts *in vivo*, we generated *Fosl2*^{Δ/Δmo} mice with a specific deletion of Fra-2 in the macrophage–osteoclast lineage (Methods). *Fosl2*^{Δ/Δmo} mice were viable and showed no apparent bone defects (Supplementary Fig. 5a, b). Despite almost complete deletion in the myeloid compartment (Supplementary Fig. 5c), osteoclasts were comparable to those in the wild type, and expression of *Bcl-2*, *HIF1 α* and HIF1 α -target genes was unaltered (Supplementary Fig. 5d). However, mutant osteoclast progenitors expressed less *LIFR* *in vitro* and showed a differentiation and fusion defect, which could be rescued by LIF addition (Supplementary Fig. 5e, f, and data not shown). This indicates that deletion of Fra-2 in the macrophage–osteoclast lineage decreases osteoclast progenitor viability *in vitro*, but not *in vivo*.

Conventional tissue culture exposes cells to 21% oxygen, which is higher than physiological levels²¹. Our results suggest that downregulation of Fra-2 or LIF/LIFR sensitizes osteoclasts to high oxygen *in vitro*. Progenitors deficient in Fra-2 or LIF expressed significantly less *Bcl-2* than controls at 21% oxygen (Fig. 3c), whereas at 3% oxygen, *Bcl-2* expression, progenitor cell survival and osteoclast differentiation of both Fra-2-deficient and LIF-deficient osteoclasts was enhanced (Fig. 3c).

Oxygen levels are sensed by prolyl hydroxylase domain (PHD) enzymes (PHD1, PHD2 and PHD3), which regulate the stability of HIF1 proteins²². PHD2 and PHD3 are induced by hypoxia in a HIF1-dependent manner²². A significant decrease in *Phd2* mRNA was observed in bones deficient in Fra-2 or LIF, whereas *Phd1* was unaffected and *Phd3* decreased only in LIF-deficient bones (Fig. 3d).

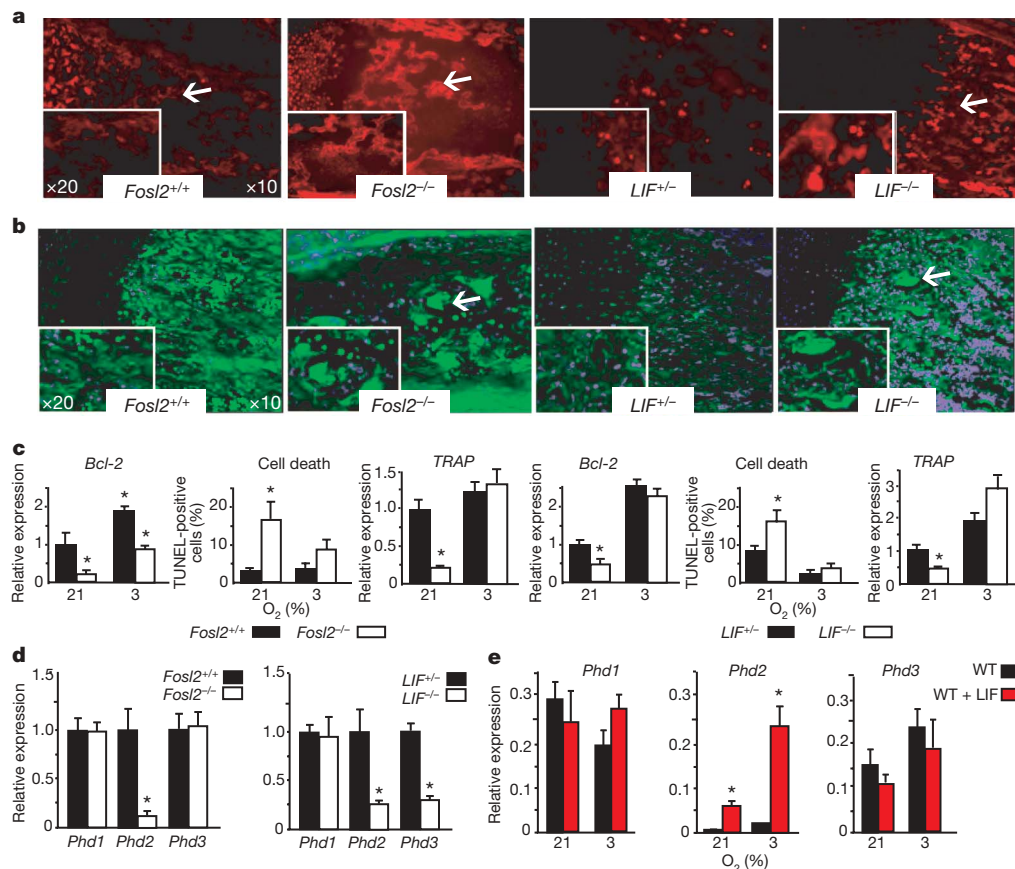


Figure 3 | Hypoxia in Fra-2-deficient and LIF-deficient long bones.

a, b, Detection of hypoxia at P1 (**a**) and HIF1 α at P2 (**b**) in *Fosl2*^{+/+}, *Fosl2*^{-/-}, *LIF*^{+/+} and *LIF*^{-/-} tibiae. Arrows indicate osteoclasts. **c**, *Bcl-2*, TRAP (qPCR) and TUNEL-positive cells in *Fosl2*^{+/+}, *Fosl2*^{-/-}, *LIF*^{+/+} and *LIF*^{-/-} fetal liver progenitors cultured 2 days (TUNEL) or 4 days (TRAP and

Bcl-2) in 21% or 3% oxygen ($n = 3$). **d**, *Phd1*, *Phd2* and *Phd3* mRNA in *Fosl2*^{+/+}, *Fosl2*^{-/-}, *LIF*^{+/+} and *LIF*^{-/-} long bones (qPCR, P2, wild type set to 1, $n = 3$). **e**, *Phd1*, *Phd2* and *Phd3* mRNA in wild-type (WT) fetal liver progenitors at 21% or 3% oxygen with or without LIF (qPCR, 48h, $n = 3$). Error bars represent s.d.; asterisk, $P < 0.01$.

Furthermore, LIF significantly induced *Phd2* mRNA expression in osteoclast cultures at both 3% and 21% oxygen, whereas *Phd1* and *Phd3* were unaffected (Fig. 3e). This suggests that *Phd2* is a downstream target of LIF/LIFR signalling regulating HIF1 α levels in bone.

PHD2 is the key oxygen sensor regulating HIF1 α under physiological conditions. Disruption of *Phd2* in mice leads to a significant increase in HIF1 α proteins and embryonic lethality, whereas *Phd1*^{-/-} and *Phd3*^{-/-} mice are viable and healthy²³. Severe placental defects

preceded the death of *Phd2*^{-/-} embryos²³. Because placentation is also disrupted in *LIFR*^{-/-} fetuses, which have giant osteoclasts, we investigated whether a placental dysfunction could contribute to the osteoclast phenotype of *Fosl2* mutants.

The expression of *LIF* and *LIFR* mRNA, but also that of *Phd2* mRNA, was significantly decreased in E18.5 *Fosl2*^{-/-} placentas in comparison with controls (Fig. 4a). Furthermore, the expression of markers for placental giant cells (*Pl1*) and spongiotrophoblasts (*Mash2*, *Tbpb*) was increased, whereas syncytiotrophoblast markers (*Tfeb*, *Gcm1*) were unaffected (Fig. 4b). HIF1 α was upregulated along with *VEGF*, *HGF*, *Bcl-2* and *Bim* expression (Fig. 4c, d and Supplementary Fig. 6a).

To address whether Fra-2 inactivation in the placenta is necessary to trigger giant hypoxic osteoclasts, *Fosl2*^{Δ/Δ} mice with specific deletion of Fra-2 in the embryo but not in the placenta were generated (Methods). *Fosl2*^{Δ/Δ} mice were viable and newborn animals showed no apparent osteoclast defects (Fig. 4e, Supplementary Fig. 6b, and data not shown). Despite efficient deletion in the bone, osteoclasts were comparable to those in the wild type, and no hypoxia was detected (Supplementary Fig. 6c, d). Furthermore, *Phd2*, *LIF* and *LIFR* were unaffected and no increased expression of HIF1 α , its downstream targets or *Bcl-2* could be detected in the long bones of *Fosl2*^{Δ/Δ} pups (Supplementary Fig. 6e–g). This indicates that inactivation of Fra-2 in the placenta is responsible for the induction of hypoxia and giant osteoclasts in the long bones of *Fosl2*^{-/-} newborn animals.

Hypoxia has a well-established physiological function in cartilage development²⁴, but its impact on osteoclast function has not been studied. Using *in vivo* and *in vitro* approaches, we have convincingly shown that genetic inactivation of the transcription factor Fra-2 or the cytokine LIF sensitized osteoclast progenitors to high oxygen and triggered hypoxia in long bones. As a consequence, osteoclasts developed in increased numbers and size in pups deficient in Fra-2 and LIF. Decreased LIF/LIFR signalling, PHD2 downregulation and increased HIF1 α expression leading to Stat-1/P-Stat-1 downregulation^{25,26}, which in turn enhances *Bcl-2* expression²⁷, are the likely causes for the increased survival and fusion of osteoclast precursors¹⁵ (Fig. 4f). Although impaired vascular oxygen supply within the bone would have been the most obvious explanation for hypoxia, our preliminary observations indicate that vascularization is unaffected in bones deficient in Fra-2 or LIF (data not shown). The placental defects in the *Fosl2* mutants together with the rescue of the hypoxia and the giant-osteoclast phenotype in the *Fosl2*^{Δ/Δ} pups show that a functional Fra-2/LIF/LIFR/PHD2/HIF axis in the placenta is essential for proper bone oxygenation (Fig. 4f). A defective placenta is probably causal of the giant osteoclasts observed in *Fosl2* and *LIF* knockout mutants but also in *LIFR* knockout mutants. *LIFR*-inactivating mutations were found in a subset of patients with Stüve–Wiedemann syndrome²⁸, a rare condition in infants characterized by low bone mass and giant osteoclasts. It is tempting to speculate that mutations in *LIF* or *Fosl2* could also be associated with this lethal disease and that hypoxia and a defective placenta could be the mediator.

Adaptation to hypoxia is a critical event in numerous pathological settings, such as tumour progression and survival of tissues in which blood flow has been suddenly interrupted. It will be of interest to find out whether inactivation of Fra-2, LIF and PHD2 in adult mice would trigger giant osteoclast formation in response to local hypoxia in bone. In contrast, upregulation of IL-6 cytokines as well as *Bcl-2* were described in Paget's disease of bone²⁹, a chronic and focal skeletal disorder characterized by an increase in both the size and number of osteoclasts. This suggests that a hypoxic response could be involved in Paget's disease and also in other destructive bone lesions such as multiple myeloma and rheumatoid arthritis. Our findings therefore establish a fundamental and unique role of Fra-2, LIF/LIFR signalling, *Bcl-2* and placenta-induced hypoxia in the regulation of osteoclast function *in vivo* and provide the basis for developing new strategies aimed at manipulating the activity and survival of

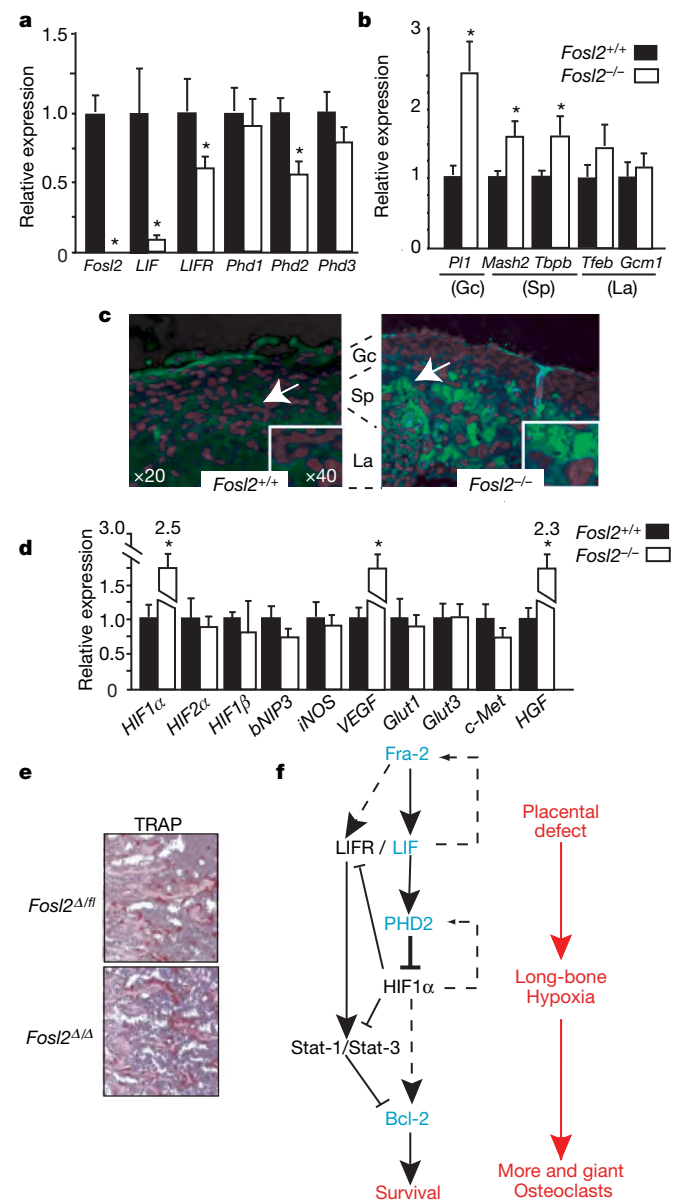


Figure 4 | A placental origin for the giant osteoclast defect. **a, b**, *Fosl2*, *LIF*, *LIFR*, *Phd1*, *Phd2* and *Phd3* (**a**) and *Pl1*, *Mash2*, *Tbpb*, *Tfeb* and *Gcm1* (**b**) mRNA in *Fosl2*^{+/+} and *Fosl2*^{-/-} E18.5 placentas (qPCR, wild type set to 1, *n* = 3). Gc, giant cells; Sp, spongiotrophoblasts; La, labyrinth (syncytiotrophoblasts). **c**, HIF1 α immunohistochemistry in *Fosl2*^{+/+} and *Fosl2*^{-/-} E18.5 placentas. Original magnifications $\times 20$ and $\times 40$ (insets). Arrows point to the areas enlarged in the insets. **d**, The relative mRNA expression of HIF and HIF-targets in *Fosl2*^{+/+} and *Fosl2*^{-/-} E18.5 placentas (qPCR, wild-type set to 1, *n* = 3). Numbers over bars represent relative expression. Error bars represent s.d.; asterisk, *P* < 0.01. **e**, TRAP staining of *Fosl2*^{Δ/Δ} and *Fosl2*^{Δ/Δ} tibiae at P2. Original magnification $\times 40$. **f**, Schematic model of how Fra-2 and LIF/LIFR signalling controls osteoclast numbers and size. Dotted lines depict hypothetical pathways, and solid lines show interactions identified in this study (proteins in blue) or in published reports. The biological phenotypes are shown in red on the right.

osteoclasts, improving on disease diagnostics and understanding the pathogenesis of bone-destructive diseases.

METHODS SUMMARY

At least three animals were used per genotype. Experiments were performed in replicates and error bars represent s.d. $P < 0.01$ (Student's *t*-test, indicated by an asterisk) was accepted as statistically significant.

Bones were embedded in paraffin (immunohistochemistry) or methylmethacrylate (histomorphometry). Trap, hypoxia and apoptosis were detected with leukocyte acid phosphatase (Sigma), hypoxypromote (Chemicon) and TUNEL (Roche) kits, respectively. Western blots were normalized to actin and quantitative PCR (qPCR) to at least one housekeeping gene. Fra-2, JunB, JunD, c-Jun, (Santa Cruz), H3K4me3 (Upstate) and H3K27me3 (a gift from T. Jenuwein) antibodies were used for ChIP. Input, bound and unbound fragments were quantified by qPCR. The proximal LIF enhancer fragment was cloned into pGL4 (Promega), transfected¹¹ in Cos7 cells and analysed with the Dual Luciferase Kit (Promega).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 6 December 2007; accepted 22 April 2008.

Published online 11 June 2008.

- Karsenty, G. & Wagner, E. F. Reaching a genetic and molecular understanding of skeletal development. *Dev. Cell* **2**, 389–406 (2002).
- Ehrlich, L. A. & Roodman, G. D. The role of immune cells and inflammatory cytokines in Paget's disease and multiple myeloma. *Immunol. Rev.* **208**, 252–266 (2005).
- Grigoriadis, A. E. *et al.* c-Fos: a key regulator of osteoclast–macrophage lineage determination and bone remodeling. *Science* **266**, 443–448 (1994).
- Karreth, F., Hoebertz, A., Scheuch, H., Eferl, R. & Wagner, E. F. The AP1 transcription factor Fra2 is required for efficient cartilage development. *Development* **131**, 5717–5725 (2004).
- Eferl, R., Zenz, R., Theussl, H. C. & Wagner, E. F. Simultaneous generation of fra-2 conditional and fra-2 knock-out mice. *Genesis* **45**, 447–451 (2007).
- Ware, C. B. *et al.* Targeted disruption of the low-affinity leukemia inhibitory factor receptor gene causes placental, skeletal, neural and metabolic defects and results in perinatal death. *Development* **121**, 1283–1299 (1995).
- Sims, N. A. *et al.* Glycoprotein 130 regulates bone turnover and bone size by distinct downstream signaling pathways. *J. Clin. Invest.* **113**, 379–389 (2004).
- Kawasaki, K. *et al.* Osteoclasts are present in gp130-deficient mice. *Endocrinology* **138**, 4959–4965 (1997).
- Hsu, L. W. & Heath, J. K. Identification of two elements involved in regulating expression of the murine leukaemia inhibitory factor gene. *Biochem. J.* **302**, 103–110 (1994).
- Martens, J. H. *et al.* The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* **24**, 800–812 (2005).
- Bakiri, L., Matsuo, K., Wisniewska, M., Wagner, E. F. & Yaniv, M. Promoter specificity and biological activity of tethered AP-1 dimers. *Mol. Cell. Biol.* **22**, 4952–4964 (2002).
- Escary, J. L., Perreau, J., Dumenil, D., Ezine, S. & Brulet, P. Leukaemia inhibitory factor is necessary for maintenance of haematopoietic stem cells and thymocyte stimulation. *Nature* **363**, 361–364 (1993).
- Stewart, C. L. *et al.* Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature* **359**, 76–79 (1992).

- Lagasse, E. & Weissman, I. L. Enforced expression of Bcl-2 in monocytes rescues macrophages and partially reverses osteopetrosis in op/op mice. *Cell* **89**, 1021–1031 (1997).
- McGill, G. G. *et al.* Bcl2 regulation by the melanocyte master regulator Mitf modulates lineage survival and melanoma cell viability. *Cell* **109**, 707–718 (2002).
- Egle, A., Harris, A. W., Bath, M. L., O'Reilly, L. & Cory, S. VavP-Bcl2 transgenic mice develop follicular lymphoma preceded by germinal center hyperplasia. *Blood* **103**, 2276–2283 (2004).
- Sasabe, E., Tatemoto, Y., Li, D., Yamamoto, T. & Osaki, T. Mechanism of HIF-1 α -dependent suppression of hypoxia-induced apoptosis in squamous cell carcinoma cells. *Cancer Sci.* **96**, 394–402 (2005).
- Arnett, T. R. *et al.* Hypoxia is a major stimulator of osteoclast formation and bone resorption. *J. Cell. Physiol.* **196**, 2–8 (2003).
- Schipani, E. *et al.* Hypoxia in cartilage: HIF-1 α is essential for chondrocyte growth arrest and survival. *Genes Dev.* **15**, 2865–2876 (2001).
- Adamopoulos, I. E., Xia, Z., Lau, Y. S. & Athanasou, N. A. Hepatocyte growth factor can substitute for M-CSF to support osteoclastogenesis. *Biochem. Biophys. Res. Commun.* **350**, 478–483 (2006).
- Utting, J. C. *et al.* Hypoxia inhibits the growth, differentiation and bone-forming capacity of rat osteoblasts. *Exp. Cell Res.* **312**, 1693–1702 (2006).
- Berra, E., Ginouves, A. & Pouyssegur, J. The hypoxia-inducible-factor hydroxylases bring fresh air into hypoxia signalling. *EMBO Rep.* **7**, 41–45 (2006).
- Takeda, K. *et al.* Placental but not heart defects are associated with elevated hypoxia-inducible factor α levels in mice lacking prolyl hydroxylase domain protein 2. *Mol. Cell. Biol.* **26**, 8336–8346 (2006).
- Schipani, E. Hypoxia and HIF-1 α in chondrogenesis. *Ann. NY Acad. Sci.* **1068**, 66–73 (2006).
- Jeong, C. H. *et al.* Hypoxia-inducible factor-1 α inhibits self-renewal of mouse embryonic stem cells *in vitro* via negative regulation of the leukemia inhibitory factor–STAT3 pathway. *J. Biol. Chem.* **282**, 13672–13679 (2007).
- Ivanov, S. V., Salnikow, K., Ivanova, A. V., Bai, L. & Lerman, M. I. Hypoxic repression of STAT1 and its downstream genes by a pVHL/HIF-1 target DEC1/STRAT1. *Oncogene* **26**, 802–812 (2007).
- Stephanou, A., Brar, B. K., Knight, R. A. & Latchman, D. S. Opposing actions of STAT-1 and STAT-3 on the Bcl-2 and Bcl-x promoters. *Cell Death Differ.* **7**, 329–330 (2000).
- Dagoneau, N. *et al.* Null leukemia inhibitory factor receptor (LIFR) mutations in Stuve–Wiedemann/Schwartz–Jampel type 2 syndrome. *Am. J. Hum. Genet.* **74**, 298–305 (2004).
- Brandwood, C. P. *et al.* Apoptotic gene expression in Paget's disease: a possible role for Bcl-2. *J. Pathol.* **201**, 504–512 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Schipani for providing protocols and reagents for hypoxia stainings; A. Schleiffer and A. Souabni for experimental advice; and A. Grigoriadis, C. Hartmann, T. Jenuwein, R. Johnson and M. Sibilia for numerous suggestions and critical reading of the manuscript. The Institute of Molecular Pathology is funded by Boehringer Ingelheim and this work was supported by the Austrian Industrial Research Promotion Fund. L.B. is partly funded by the NoE Cells into Organs (LSHM-CT-2003-504468). A.H. was the recipient of an EMBO and Marie Curie long-term fellowship. M.A. is funded by the German Research Community (DFG).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to E.F.W. (wagner@imp.ac.at or ewagner@cno.es).

METHODS

Mice and serum measurements. *Fosl2*^{Δ/Amo} mice were generated by crossing mice with floxed *Fosl2* alleles to mice carrying a *LysM-Cre* allele^{4,30}. *Fosl2*^{Δ/Δ} mice were generated by crossing mice with floxed *Fosl2* alleles to mice carrying a *MORE-Cre* allele with specific Cre expression in the embryo but not in the placenta³¹. *Fosl2*^{+/-}, *Fosl2*^{Δ/Amo}, *Vav2-hBcl-2* (ref. 16), *Fosl2*^{fl/+}; *MORE-Cre* (ref. 31) and *LIF*^{-/-} (ref. 13) mice were maintained in C57BL/6x129sv, C57BL/6 (*Vav2-hBcl-2*) and MF1 (*LIF*^{-/-}) backgrounds. Genotyping primers are available from the authors on request. Deoxyypyridinoline crosslinks were measured in serum from newborn animals and normalized to creatinine with commercially available kits (Metra Biosystems).

Histological analysis and histomorphometry. Limbs were fixed overnight in 3.7% formaldehyde in PBS. For histological, immunohistochemical and *in situ* hybridization analyses, bones were dehydrated and embedded in paraffin; 5-μm sections were used. For histomorphometry, undecalcified tibiae and lumbar spines were embedded in methylmethacrylate and 5-μm sections were stained with toluidine blue and modified von Kossa/van Gieson. Quantitative histomorphometry was performed in a blinded fashion in accordance with standardized protocols³² using the Osteomeasure histomorphometry system (Osteometrix).

TRAP and hypoxia staining, immunohistochemistry and *in situ* hybridization. TRAP staining was performed on deparaffinized sections with the leukocyte acid phosphatase kit (Sigma) or ELF97 (Molecular Probes³³). Hypoxia staining¹⁹ was performed as described with a hypoxyprobe kit (Chemicon). Immunohistochemistry for LIF (R&D), LIFR (Santa Cruz), P-Stat-1 (R&D) and Bcl-2 (Santa Cruz) were performed on deparaffinized sections with an automation immunohistochemistry Discovery XT system (Ventana Medical Systems). For *in situ* hybridization⁴, digoxigenin-labelled riboprobes were synthesized with a DIG RNA labelling kit and the signal was detected with BM-purple AP substrate (Roche).

Osteoclast cultures. Fetal livers were harvested from E14.5 embryos, brought into suspension and cultured in 24-well plates (5×10^5 cells) with macrophage colony-stimulating factor (20 ng ml^{-1} ; R&D) and RANKL (5 ng ml^{-1} ; R&D) and cultured for 2–7 days. Bone marrow cultures were performed in a similar way. LIF (R&D) was used at 10 pg ml^{-1} . All cell culture assays were performed at least twice with four to six replicates per experiment.

RNA isolation and reverse transcription. Total RNA was isolated with TRIzol (Invitrogen). cDNA synthesis was performed using $2 \mu\text{g}$ RNA with the Ready-To-Go You-Prime-It First-Strand Beads and random primers (Amersham Biosciences).

Real-time PCR. qPCR reactions were performed with SYBR Green (Molecular Probes) and an Opticon2 Monitor Fluorescence Thermocycler (MJ Research). Primer sequences are available upon request. The comparative Ct (cycle threshold) method was used to quantify the amplified fragments. RNA expression levels were normalized to at least one housekeeping gene (encoding tubulin,

actin, glyceraldehyde-3-phosphate dehydrogenase or hypoxanthine–guanine phosphoribosyltransferase). Values are presented as relative expression and wild-type levels are set to 1. Chromatin-bound fragments were normalized to the same fragments amplified from input chromatin.

Western blotting. Whole long-bone extracts were prepared with cell lysis buffer³⁴ and western blots were performed with the following antibodies: MITF (Lab vision), gp130 (Santa Cruz), LIF (R&D), LIFR (Santa Cruz), P-Stat-1 (R&D), P-Stat-3 (R&D), ERK (Santa Cruz), P-ERK-1 and P-ERK-2 (Santa Cruz), p38, P-p38, P-AKT (Cell Signalling), Bcl-2 (Santa Cruz), Mcl-1 (Abcam), HIF1α (R&D) and iNOS (Santa Cruz). Blots were quantified with ImageJ software. Values were normalized to actin.

Chromatin immunoprecipitation (ChIP). ChIP¹⁰ was performed with antibodies against Fra-2, JunB, JunD, c-Jun (Santa Cruz) H3K4me3 (Upstate) and H3K27me3 (a gift from T. Jenuwein). The specific binding of each Jun antibody was verified with a corresponding pair of wild-type and knockout osteoblasts, and the *Fosl1* canonical TRE element as a target sequence. Primer pairs spanning 4 kilobases 5' of the *LIF* gene were first used to locate the Fra-2 binding regions in the LIF promoter. Bound and unbound fragments were quantified by qPCR.

Reporter assays. The *LIF* promoter reporter was constructed by cloning a PCR-amplified fragment (−1,900 to −1,550 base pairs upstream of the start site) containing the proximal TRE-like element, into the pGL4 luciferase reporter (Promega). The TRE site was mutated with mutAgene kit (Stratagene). Cos7 cells (6×10^4 per well) were plated in 24-well dishes. Luciferase reporter construct ($1.5 \mu\text{g}$), $0.2 \mu\text{g}$ of the renilla internal control (pHRG-tk Promega) and $0.2\text{--}1 \mu\text{g}$ of each AP-1 expression vector were co-transfected in triplicate with the use of calcium phosphate¹¹. Luciferase activity was quantified with the Dual Luciferase Kit (Promega) with a Szabo Scandic luminometer.

Statistical analysis. All experiments were repeated at least twice and performed in triplicate. Statistical analysis was performed with Student's *t*-test; $P < 0.01$ was accepted as significant. Data are shown as means and the error bars represent s.d.

30. Clausen, B. E., Burkhardt, C., Reith, W., Renkawitz, R. & Forster, I. Conditional gene targeting in macrophages and granulocytes using *LysMcre* mice. *Transgenic Res.* **8**, 265–277 (1999).
31. Tallquist, M. D. & Soriano, P. Epiblast-restricted Cre expression in MORE mice: a tool to distinguish embryonic vs. extra-embryonic gene function. *Genesis* **26**, 113–115 (2000).
32. Parfitt, A. M. *et al.* Bone histomorphometry: standardization of nomenclature, symbols, and units. Report of the ASBMR Histomorphometry Nomenclature Committee. *J. Bone Miner. Res.* **2**, 595–610 (1987).
33. Filgueira, L. Fluorescence-based staining for tartrate-resistant acid phosphatase (TRAP) in osteoclasts combined with other fluorescent dyes and protocols. *J. Histochem. Cytochem.* **52**, 411–414 (2004).
34. Bakiri, L. *et al.* Role of heterodimerization of c-Fos and Fra1 proteins in osteoclast differentiation. *Bone* **40**, 867–875 (2007).

LETTERS

IRF4 addiction in multiple myeloma

Arthur L. Shaffer¹, N. C. Tolga Emre¹, Laurence Lamy¹, Vu N. Ngo¹, George Wright², Wenming Xiao³, John Powell³, Sandeep Dave¹, Xin Yu¹, Hong Zhao¹, Yuxin Zeng⁴, Bangzheng Chen⁴, Joshua Epstein⁴ & Louis M. Staudt¹

The transcription factor IRF4 (interferon regulatory factor 4) is required during an immune response for lymphocyte activation and the generation of immunoglobulin-secreting plasma cells^{1–3}. Multiple myeloma, a malignancy of plasma cells, has a complex molecular aetiology with several subgroups defined by gene expression profiling and recurrent chromosomal translocations^{4,5}. Moreover, the malignant clone can sustain multiple oncogenic lesions, accumulating genetic damage as the disease progresses^{6,7}. Current therapies for myeloma can extend survival but are not curative^{8,9}. Hence, new therapeutic strategies are needed that target molecular pathways shared by all subtypes of myeloma. Here we show, using a loss-of-function, RNA-interference-based genetic screen, that IRF4 inhibition is toxic to myeloma cell lines, regardless of transforming oncogenic mechanism. Gene expression profiling and genome-wide chromatin immunoprecipitation analysis uncovered an extensive network of IRF4 target genes and identified *MYC* as a direct target of IRF4 in activated B cells and myeloma. Unexpectedly, *IRF4* was itself a direct target of *MYC* transactivation, generating an autoregulatory circuit in myeloma cells. Although IRF4 is not genetically altered in most myelomas, they are nonetheless addicted to an aberrant IRF4 regulatory network that fuses the gene expression programmes of normal plasma cells and activated B cells.

Recently, we developed a genetic method to identify therapeutic targets in cancer in which small hairpin RNAs (shRNAs) that mediate RNA interference are screened for their ability to block cancer cell proliferation and/or survival¹⁰. Here we report the results of such an ‘Achilles’ heel’ screen in multiple myeloma (Supplementary Table 3). We used myeloma cell lines from three molecular subtypes: KMS12 (*CCND1* translocation), H929 (*FGFR3-MMSET* translocation) and SKMM1 (*MAFB*, *IRF4* translocations). Myeloma cells that received an shRNA targeting the coding region of *IRF4* were depleted from cultures by 2–8-fold (Fig. 1a). Lymphoma cell lines were largely unaffected by IRF4 knockdown, with the exception of OCI-Ly3, an activated B-cell-like diffuse large B cell lymphoma line that expresses IRF4 highly¹¹. We next identified two further shRNAs against *IRF4* that were toxic to myeloma cell lines, one of which was directed against the *IRF4* 3′ untranslated region (UTR; Supplementary Fig. 1). The toxicity of this shRNA was associated with a 50%–75% decrease in IRF4 messenger RNA and protein (Supplementary Fig. 2a–c). Cell death occurred within 3 days, as measured by an increase in sub-G1 DNA content; there was, however, no effect on the cell cycle (Supplementary Fig. 2d–g). Expression of a complementary DNA containing only the coding region of *IRF4* was able to rescue myeloma cells from the toxicity of the 3′ UTR-directed *IRF4* shRNA, confirming that the toxicity of this shRNA was specific (Fig. 1b).

The knockdown of IRF4 killed ten myeloma cell lines, but had a minimal effect on five lymphoma cell lines (Fig. 1c). These myeloma lines bear many of the recurrent genetic aberrations typical of this

cancer, including translocations of *CCND1*, *MYC*, *MAF*, *MAFB*, *FGFR3-MMSET*, *NIK* and *IRF4*, as well as *RAS* mutations, inactivation of *TP53* and *CDKN2C*, and genetic abnormalities that activate the NF-κB pathway (Supplementary Table 1). Resequencing of the *IRF4* coding exons in these lines showed that nine had a wild-type sequence and one had a heterozygous mutation in exon 8 resulting in a missense substitution, the functional significance of which is unknown. Moreover, no amplification of the *IRF4* locus was detected by array-based comparative genomic hybridization and no translocations involving *IRF4* were detected by cytogenetics, with the exception of the previously documented *IRF4* translocation in SKMM1 cells (data not shown). Thus, IRF4 dependency spans many myeloma

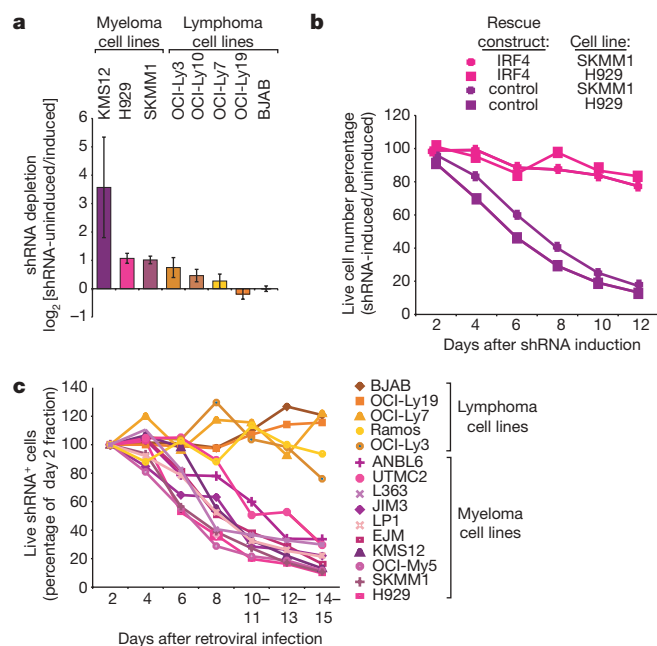


Figure 1 | IRF4 is required for myeloma cell survival. **a**, Cell lines were screened using a retrovirally delivered, doxycycline-inducible, shRNA library to identify genes required for cell growth or survival, as described¹⁰. Depletion of cells bearing an *IRF4*-targeted shRNA in shRNA-induced versus uninduced cells is plotted; error bars represent the s.d. of triplicate measurements. **b**, Expression of the *IRF4* coding region rescues myeloma cells from lethality of an shRNA targeting the *IRF4* 3′ UTR (see text for details). **c**, An *IRF4* shRNA is toxic to myeloma but not lymphoma cell lines. A vector for constitutive expression of *IRF4* shRNA was transduced into cell lines, and viability of shRNA⁺ cells was monitored. In **b** and **c**, cells expressing shRNA were monitored by flow cytometry for a co-expressed GFP marker and data were normalized to the percentage of GFP⁺ cells at day 2 after infection.

¹Metabolism Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. ²Biometric Research Branch, National Cancer Institute, Rockville, Maryland 20892, USA. ³Bioinformatics and Molecular Analysis Section, Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892, USA. ⁴Myeloma Institute for Research and Therapy, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72212, USA.

subtypes and does not require genetic abnormalities in the *IRF4* locus.

To understand the molecular basis for this dependency, we identified downstream targets of IRF4 by profiling gene expression changes in myeloma lines after induction of *IRF4* shRNAs (Supplementary Fig. 3). A total of 308 genes were consistently down-regulated after IRF4 knockdown (Supplementary Table 2). This list was significantly enriched for genes that are more highly expressed in primary myeloma samples than in normal, mature B cells, on the basis of gene set enrichment analysis¹² of published gene expression profiling data ($P = 0.002$; Fig. 2a)¹³. Thus, IRF4 directs a broad gene expression programme that is characteristic of primary myeloma cells.

We next investigated whether the IRF4 target genes in myeloma are also upregulated in other normal haematopoietic cells that require high IRF4 expression, including plasma cells³, mitogenically activated mature B cells¹ and dendritic cells¹⁴. Human bone-marrow-derived plasma cells expressed 22% of the IRF4 target genes at higher levels than mature blood B cells (Fig. 2a)¹³. Similarly, 25% of the IRF4 targets were more highly expressed in dendritic cells than in monocytes (Supplementary Fig. 4)¹⁵. Blood B cells activated by anti-IgM crosslinking expressed one-third of the IRF4 target genes more highly than resting B cells (Fig. 2a).

However, IRF4 regulates a broader set of genes in myeloma than in individual haematopoietic subsets. Roughly one-quarter of the IRF4 target genes in myeloma were upregulated in activated B cells but not plasma cells, including genes known to be important in cellular growth and proliferation, such as *MYC* (Fig. 2a). Conversely, one-sixth of the myeloma IRF4 target genes were highly expressed in plasma cells but not activated B cells.

To identify direct IRF4 targets, we performed genome-wide chromatin immunoprecipitation using DNA microarrays (ChIP–CHIP) with probes spanning ~10 kb at the 5' end of 17,574 human genes.

Specific IRF4 binding to 558 genes was detected in a myeloma cell line (KMS12) but not a lymphoma cell line (OCI-Ly19). Of these, 35 were also IRF4 targets by gene expression profiling, a highly significant overlap ($P = 1.0 \times 10^{-16}$, chi-squared; Fig. 2b and Supplementary Fig. 5), and were considered presumptive direct IRF4 targets (Supplementary Table 2). Direct binding of IRF4 was confirmed by conventional chromatin immunoprecipitation for 22 genes, leading us to conclude that all 35 genes are probably IRF4 direct targets (Fig. 2b and data not shown). This list of IRF4 direct targets is a conservative estimate because the ChIP–CHIP arrays interrogate limited regions around each gene. Indeed, direct ChIP experiments demonstrated that two other IRF4 target genes, *PRDM1* and *SQLE*, were directly bound by IRF4 in regions not covered by our ChIP–CHIP analyses (Supplementary Fig. 5). IRF4 bound to the promoter and fourth intron of *PRDM1*, which encodes BLIMP-1, another important regulator of plasmacytic differentiation (Supplementary Fig. 5). These observations support the proposal that *IRF4* lies genetically upstream of *PRDM1* in the regulatory hierarchy of terminal B cell differentiation³. Notably, IRF4 bound to its own promoter, supporting a positive feedback mechanism by which plasma cells can maintain high IRF4 expression³ (Supplementary Fig. 5).

A direct IRF4 target of particular interest is *MYC* given its prominent role in the pathogenesis of myeloma¹⁶. Knockdown of *IRF4* reduced *MYC* mRNA levels by more than twofold in myeloma cell lines and caused *MYC* DNA-binding activity to decrease in nuclear extracts of myeloma cells. Conversely, ectopic expression of IRF4 in a lymphoma cell line increased *MYC* mRNA levels (Fig. 3a and Supplementary Fig. 6). Using ChIP, we surveyed regions of the *MYC* locus for binding by IRF4 in myeloma cells and detected a peak of binding around -1.6 kb upstream of the *MYC* start site, coinciding with a region detected by ChIP–CHIP (Fig. 3b and Supplementary Fig. 7). Knockdown of *IRF4* expression diminished the amount of IRF4 bound to this region of the *MYC* promoter

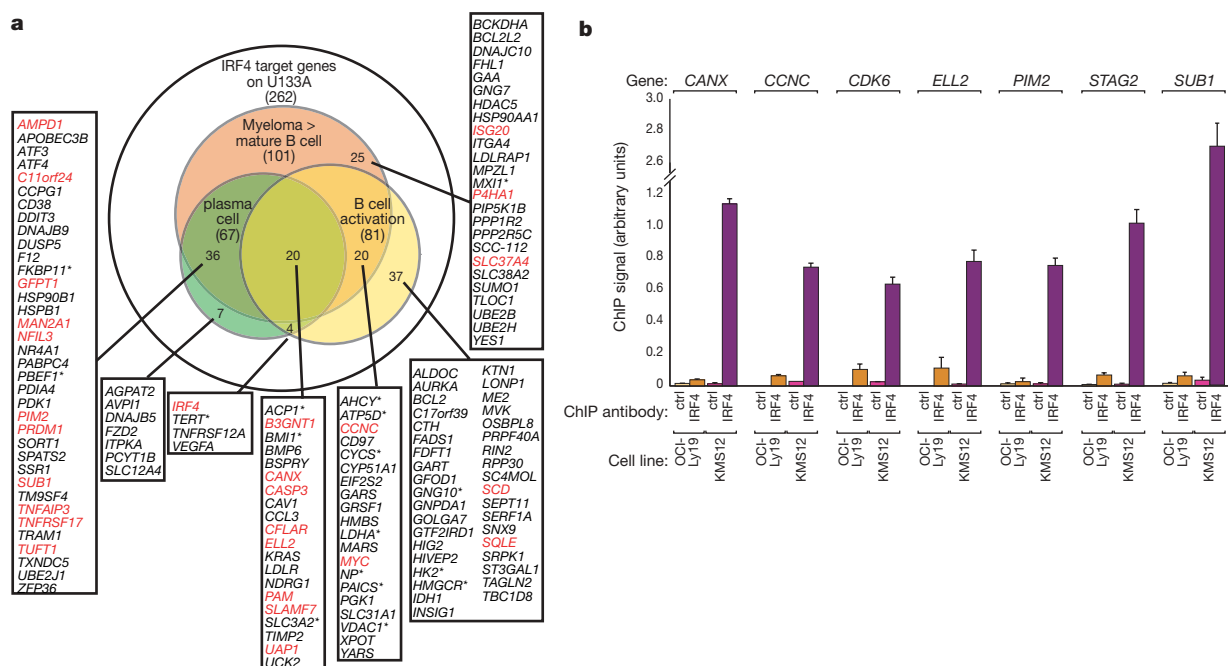


Figure 2 | IRF4 target genes in multiple myeloma. **a**, Venn diagram depicting IRF4 target genes and the overlap between the myeloma, plasma cell and activated B cell gene expression programmes. Of the 308 IRF4 target genes (Supplementary Fig. 3), 262 were well-measured on Affymetrix U133A gene expression arrays. One-hundred-and-one target genes were more highly expressed in primary myeloma samples than primary mature B cells (>1.4-fold, red circle), 67 were more highly expressed in primary plasma cells than mature B cells (>1.4-fold, green circle), and 81 were induced between 1 h and 24 h after activation of primary human B cells by anti-IgM

crosslinking (>2-fold, yellow circle). Red denotes direct IRF4 targets by ChIP; asterisk denotes direct MYC targets. **b**, Representative conventional ChIP assays for genes identified as IRF4 targets by both gene expression profiling and ChIP–CHIP. Individual ChIP assays were performed on chromatin from the KMS12 myeloma line and the OCI-Ly19 lymphoma line using either anti-IRF4 or control (ctrl) antibodies. The ChIP signal is given in arbitrary relative units calculated from quantitative PCR data, on the basis of the abundance of the indicated gene in the immunoprecipitated DNA versus input DNA. Error bars are s.d. from triplicate measurements.

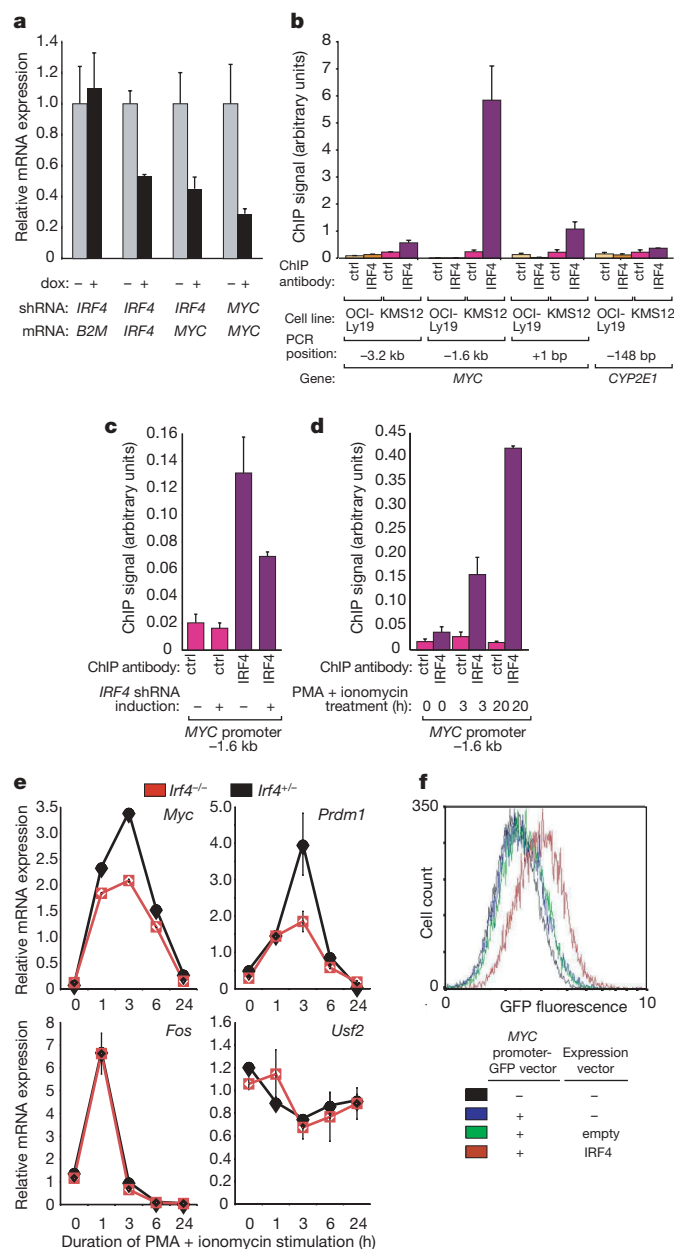


Figure 3 | MYC is a direct IRF4 target gene in myeloma and activated B cells. **a**, Knockdown of IRF4 decreases *MYC* mRNA expression. The SKMM1 myeloma line was transduced with *IRF4* or *MYC* shRNAs, and gene expression was measured by quantitative RT-PCR after 4 days of shRNA induction, normalized to the signal from uninduced cells. **b**, Binding of IRF4 to the *MYC* promoter. ChIP was performed as in Fig. 2, comparing the myeloma line KMS12 to the lymphoma line OCI-Ly19, for regions of the *MYC* promoter (as indicated relative to the transcriptional start site) or a control locus, *CYP2E1*. **c**, IRF4 knockdown decreases IRF4 binding to the *MYC* promoter. ChIP was performed using KMS12 cells with or without shRNA induction for 4 days. **d**, Activation of human blood B cells leads to IRF4 binding to the *MYC* promoter. ChIP assays were performed using purified peripheral human blood B cells, either unstimulated or activated with PMA plus ionomycin for the indicated times. **e**, Genetic deficiency of *Irf4* impairs *Myc* induction during lymphocyte activation. Quantitative RT-PCR was performed on RNA from resting splenic B cells of IRF4-deficient or wild-type mice, either unstimulated or activated with PMA and ionomycin for the indicated times. Values were normalized to *B2M* expression. **f**, IRF4 transactivates the *MYC* promoter. The OCI-Ly7 lymphoma line was transiently transfected with a GFP expression vector driven by the human *MYC* promoter alone, with an IRF4 expression vector, or with an empty vector control. Flow cytometry for GFP fluorescence is shown. Error bars indicate s.d. of triplicate measurements.

(Fig. 3c). In human B cells, phorbol myristate acetate (PMA) and ionomycin treatment induces transcription of *IRF4* and *MYC* (data not shown). Correspondingly, a sharp increase in IRF4 binding to the *MYC* promoter was detected after 3 and 20 h of PMA and ionomycin activation (Fig. 3d). Genetic evidence that *Myc* is an *Irf4* target was provided by analysis of mitogenically stimulated wild-type and *Irf4*-deficient mouse B cells (Fig. 3e). In *Irf4*-deficient cells, both *Myc* and *Prdm1* failed to be fully induced by PMA and ionomycin treatment whereas the immediate early gene *Fos* was normally induced, and a housekeeping gene, *Usf2*, did not change in expression. Finally, ectopic expression of IRF4 in a lymphoma cell line was able to transactivate a reporter construct in which GFP is under the control of the *MYC* promoter (Fig. 3f).

These data provide strong evidence implicating *MYC* as a direct target gene of IRF4. Accordingly, the list of IRF4 targets was highly enriched for genes that are directly transactivated by *MYC*^{17–19} ($n = 23$; $P = 1 \times 10^{-8}$, chi-squared; Supplementary Table 2 and Supplementary Fig. 9). These genes encode principal components of glycolysis (LDHA, HK2 and PDK1) and mitochondrial respiration (ATP5D, CYCS), as well as important regulators of cellular senescence (BMI1, TERT). As *MYC* is a key coordinator of cellular growth, metabolism and proliferation²⁰, we examined whether knockdown of *MYC* expression was toxic to myeloma cells. An shRNA targeting the *MYC* 3' UTR knocked down *MYC* expression and DNA binding by ~twofold (Supplementary Fig. 6). This shRNA was toxic to both myeloma and lymphoma cell lines but had little effect on the myeloma cell line U266, consistent with its high expression of *MYC1* instead of *MYC* (Fig. 4a)²¹. Expression of the *MYC* coding region was able to rescue cells from the toxicity of the 3' UTR-directed *MYC* shRNA, confirming its specificity (Fig. 4b). Thus, loss of *MYC* expression may contribute to the toxicity of *IRF4* shRNAs for myeloma cells.

Using two independent *MYC* shRNAs, we identified the targets of *MYC* in myeloma cells. After *MYC* shRNA induction, the expression levels of many direct *MYC* targets decreased (Fig. 4c). The expression of *IRF4* also decreased, as did the expression of many IRF4 target genes (Fig. 4c, d). ChIP demonstrated binding of *MYC* to a region of the *IRF4* first intron in two myeloma cell lines expressing *MYC* (KMS12 and H929) but not in a cell line with very low *MYC* expression (U266, Fig. 4e). Furthermore, we detected *MYC* binding to *IRF4* in mitogenically activated B cells, which express *MYC*, but not in resting B cells, which do not (Fig. 4f).

These data show a positive regulatory loop in myeloma cells in which IRF4 and *MYC* mutually reinforce the expression of each other (Fig. 5a). In keeping with this model, myeloma patient samples express both *MYC* and *IRF4* mRNA more highly than normal plasma cells ($P = 5.1 \times 10^{-7}$ for IRF4; Fig. 4g). Moreover, *MYC* and *IRF4* mRNA levels showed significant positive correlation across 451 primary myeloma samples⁴ ($r = 0.24$, $P = 2.5 \times 10^{-7}$, Supplementary Fig. 7). This moderate correlation was notable because IRF4 is probably only one of many factors regulating *MYC* transcription in myelomas²². Although the *MYC* locus in myeloma is often amplified and inserted at ectopic genomic locations, especially within and near the immunoglobulin loci¹⁶, the *MYC* breakpoints in these chromosomal rearrangements are many kilobases from the *MYC* transcriptional start site and thus preserve the IRF4 binding region. Our data suggest that the oncogenic activation of *MYC* in myeloma upregulates IRF4, which in turn drives expression of *MYC* and other IRF4 target genes (Fig. 5a).

In some respects, the dependency of myeloma cells on IRF4 is reminiscent of the function of 'lineage survival' oncogenes²³. These genes are primarily transcription factors that provide essential functions in a particular cellular lineage but are also dysregulated in cancers derived from that lineage. IRF4 differs from lineage survival oncogenes in two respects. First, many lineage survival oncogenes are altered by mutations or chromosomal structural alterations whereas the *IRF4* locus seems to be unaltered in most myelomas. Second, the

regulatory network that IRF4 controls in myeloma is decidedly abnormal, not merely reflecting the genetic programme of normal plasma cells but also borrowing from the genetic programme of antigen-stimulated mature B cells (Figs 2a and 5b). This transcriptional promiscuity is exemplified by the direct IRF4 targets *MYC*, *SCD*, *SQLE*, *CCNC* and *CDK6*, which are not highly expressed in normal plasma cells but are upregulated in mature B cells on antigen receptor signalling (Figs 2a and 5b). Thus, myelomas have broadened the genetic repertoire of IRF4, perhaps due to epigenetic alterations that allow IRF4 access to loci that are normally silenced in plasma cells. Hence, the dependency of myeloma on IRF4 may be best described as 'non-oncogene addiction'; that is, the aberrant function of a normal cellular protein that is required for cancer cell proliferation or survival²⁴.

The direct targets of IRF4 show that it is a master regulator influencing metabolic control, membrane biogenesis, cell cycle progression, cell death, transcriptional regulation and plasmacytic differentiation (Fig. 5b). Given this pleiotropic programme, we believe that loss of IRF4 in a myeloma cell results in 'death by a thousand cuts'. For

example, several important cell-cycle regulators are IRF4 targets, in keeping with its role in lymphocyte activation¹, including *STAG2*, *CDK6* and *MYC*. *STAG2* encodes a component of the cohesin complex crucially involved in the segregation of chromosomes during mitosis²⁵. Two different shRNAs targeting *STAG2* were toxic for both a myeloma and a lymphoma cell line (Supplementary Fig. 8), as were shRNAs targeting *MYC* (Fig. 4a). Similarly, myeloma cells were specifically killed by two different shRNAs targeting *SUB1*, an IRF4 direct target which encodes a transcriptional coactivator²⁶. It seems probable, therefore, that decreased expression of each of these IRF4 direct targets contributes to *IRF4* shRNA toxicity. A prominent role for IRF4 in regulating membrane biogenesis was indicated by the many enzymes and regulators of sterol and lipid synthesis under its control (Supplementary Fig. 9), including *SQLE* and *SCD*, which encode rate-limiting enzymes in these pathways. Furthermore, the IRF4 target gene set was strikingly enriched for genes encoding components of glucose metabolism and ATP production, many of which are targets of *MYC* (Supplementary Fig. 9). As a result it is plausible that metabolic collapse also contributes to cell death caused by IRF4 deprivation.

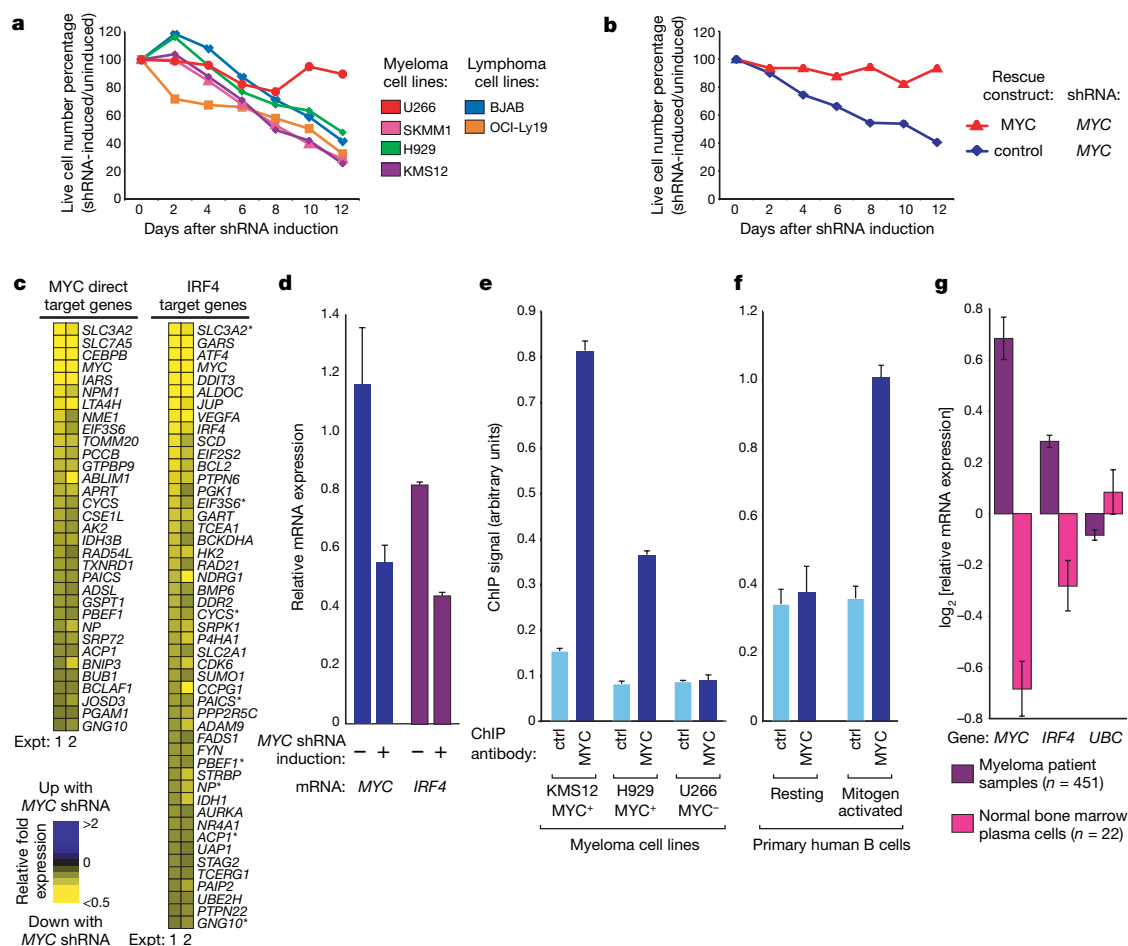


Figure 4 | *IRF4* is a direct *MYC* target gene in myeloma and activated B cells. **a**, Lethality of a *MYC* shRNA for cell lines expressing *MYC*. Cell lines were transfected with a *MYC* shRNA vector and the fraction of shRNA⁺ (GFP⁺) cells was monitored over time. All lines express *MYC*, except U266, which expresses *MYCL1*. **b**, Expression of the *MYC* coding region rescues H929 myeloma cells from lethality of an shRNA targeting the *MYC* 3' UTR. **c**, *MYC* knockdown downregulates *MYC* direct target genes and *IRF4* target genes. KMS12 myeloma cells were induced for *MYC* shRNA expression for 4 days and profiled for gene expression changes. Each experiment used a different *MYC* shRNA. Exemplar array elements are shown (reduced by >1.3-fold in both experiments) for known *MYC* direct targets¹⁷ and *IRF4* targets (this work). **d**, *MYC* knockdown decreases *IRF4* mRNA expression. Shown are quantitative RT-PCR measurements of *MYC* and *IRF4* mRNA

levels in KMS12 myeloma cells, with or without induction of *MYC* shRNA. Error bars indicate s.d. of triplicate measurements. **e**, *MYC* binds to the *IRF4* locus. ChIP of *MYC* binding to the *IRF4* first intron in myeloma cells expressing *MYC* (KMS12 and H929), but not in the myeloma line U266 that lacks high *MYC* expression. **f**, *MYC* binding to the *IRF4* locus is induced in activated primary human B cells. ChIP of *MYC* binding to the *IRF4* first intron in human blood B cells, either unstimulated or activated with PMA and ionomycin for 6 h. ctrl, control. **g**, *MYC* and *IRF4* are more highly expressed in primary myeloma patient samples than in normal human bone marrow plasma cells. Previously published gene expression profiling data⁴ was analysed for average mRNA expression of *MYC*, *IRF4* and a control gene, *UBC*.

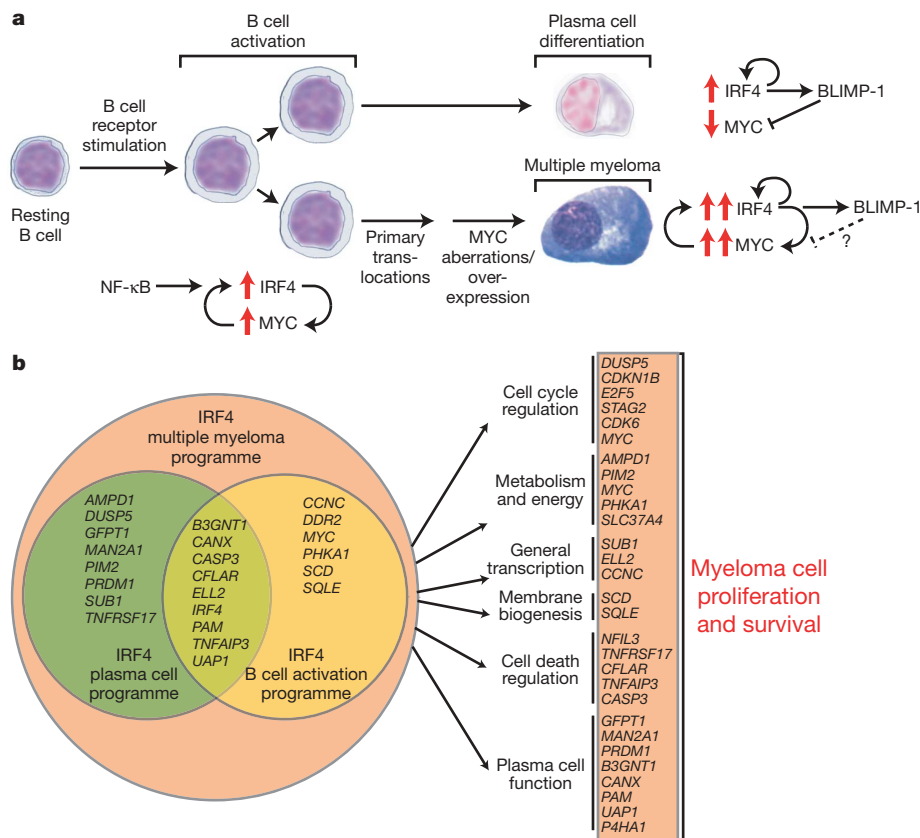


Figure 5 | Model of IRF4 control over B cell development and multiple myeloma oncogenesis. **a**, IRF4 and MYC form a positive autoregulatory loop during normal B cell activation and in multiple myeloma. Genetic abnormalities of MYC upregulate its expression in myeloma, thereby augmenting IRF4 expression. In normal plasma cells, BLIMP-1 represses MYC, but this control circuit may be abrogated in myeloma. **b**, IRF4 as a

master regulator of the myeloma phenotype. IRF4 controls a myeloma-specific gene expression programme that fuses the IRF4 regulatory programmes from activated B cells and plasma cells. IRF4 direct targets regulate many essential cellular processes, causing myeloma cells to be addicted to IRF4.

Our data demonstrate that myelomas are addicted to an abnormal regulatory network controlled by IRF4, irrespective of their molecular subtype and underlying oncogenic abnormalities. Hence, IRF4 emerges as a master regulator of an aberrant and malignancy-specific gene expression programme relevant to all molecular subtypes of this cancer. Because mice lacking one allele of *Irf4* are phenotypically normal¹ and as a ~50% knockdown of *IRF4* mRNA and protein was sufficient to kill myeloma cell lines, a therapeutic window may exist in which IRF4-directed therapy might kill myeloma cells while sparing normal cells. Although transcription factors have been considered intractable therapeutic targets, recent successful targeting of p53 (ref. 27) and BCL-6 (ref. 28) provides hope that IRF4 can be exploited as an Achilles' heel of multiple myeloma.

METHODS SUMMARY

Lines were engineered to express the ecotropic retroviral receptor and the bacterial tetracycline repressor as described¹⁰. The retroviral constructs for shRNA expression and the design of shRNA library sequences have been described¹⁰; in some vectors, the puromycin selectable marker (puro) was replaced by a fusion between puro and green fluorescent protein (GFP) for tracking by flow cytometry. Doxycycline (20 ng ml⁻¹) was used for shRNA induction. IRF4 and MYC were expressed using retroviral vectors as described³. Primary human resting blood B cells were purified by magnetic separation (CD19⁺ beads, Miltenyi) and grown at 2 million cells ml⁻¹ in IMDM containing 10% fetal bovine serum; primary mouse splenic, resting B cells were purified by magnetic separation (B cell kit, Miltenyi) and grown at 2 million cells ml⁻¹ in RPMI plus 10% fetal bovine serum. Lymphocytes were activated with PMA (40 ng ml⁻¹) and ionomycin (2 μM). Gene expression profiling was performed using Agilent 4 × 44K or Lymphochip²⁹ microarrays. ChIP-CHIP experiments were performed using Agilent Human Promoter Set microarrays.

Received 20 March; accepted 9 May 2008.
Published online 22 June 2008.

- Mittrucker, H. W. *et al.* Requirement for the transcription factor LSIRF/IRF4 for mature B and T lymphocyte function. *Science* **275**, 540–543 (1997).
- Klein, U. *et al.* Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nature Immunol.* **7**, 773–782 (2006).
- Sciammas, R. *et al.* Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. *Immunity* **25**, 225–236 (2006).
- Zhan, F. *et al.* The molecular classification of multiple myeloma. *Blood* **108**, 2020–2028 (2006).
- Bergsagel, P. L. *et al.* Promiscuous translocations into immunoglobulin heavy chain switch regions in multiple myeloma. *Proc. Natl Acad. Sci. USA* **93**, 13931–13936 (1996).
- Kuehl, W. M. & Bergsagel, P. L. Multiple myeloma: evolving genetic events and host interactions. *Nature Rev. Cancer* **2**, 175–187 (2002).
- Carrasco, D. R. *et al.* High-resolution genomic profiles define distinct clinicopathogenetic subgroups of multiple myeloma patients. *Cancer Cell* **9**, 313–325 (2006).
- Barlogie, B. *et al.* Treatment of multiple myeloma. *Blood* **103**, 20–32 (2004).
- Mitsiades, C. S., Mitsiades, N., Munshi, N. C. & Anderson, K. C. Focus on multiple myeloma. *Cancer Cell* **6**, 439–444 (2004).
- Ngo, V. N. *et al.* A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* **441**, 106–110 (2006).
- Davis, R. E., Brown, K. D., Siebenlist, U. & Staudt, L. M. Constitutive nuclear factor κB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. *J. Exp. Med.* **194**, 1861–1874 (2001).
- Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
- Gutierrez, N. C. *et al.* Gene expression profiling of B lymphocytes and plasma cells from Waldenström's macroglobulinemia: comparison with expression patterns of the same cell counterparts from chronic lymphocytic leukemia, multiple myeloma and normal individuals. *Leukemia* **21**, 541–549 (2007).
- Tamura, T. *et al.* IFN regulatory factor-4 and -8 govern dendritic cell subset development and their functional diversity. *J. Immunol.* **174**, 2573–2581 (2005).

15. Su, A. I. *et al.* A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl Acad. Sci. USA* **101**, 6062–6067 (2004).
16. Shou, Y. *et al.* Diverse karyotypic abnormalities of the *c-myc* locus associated with *c-myc* dysregulation and tumor progression in multiple myeloma. *Proc. Natl Acad. Sci. USA* **97**, 228–233 (2000).
17. Zeller, K. I. *et al.* Global mapping of *c-Myc* binding sites and target gene networks in human B cells. *Proc. Natl Acad. Sci. USA* **103**, 17834–17839 (2006).
18. Guney, I., Wu, S. & Sedivy, J. M. Reduced *c-Myc* signaling triggers telomere-independent senescence by regulating Bmi-1 and p16^{INK4a}. *Proc. Natl Acad. Sci. USA* **103**, 3645–3650 (2006).
19. Kim, J. W., Gao, P., Liu, Y. C., Semenza, G. L. & Dang, C. V. HIF-1 and dysregulated *c-Myc* cooperatively induces VEGF and metabolic switches, HK2 and PDK1. *Mol Cell Biol* **27**, 7381–7393 (2007).
20. Dang, C. V. *et al.* The *c-Myc* target gene network. *Semin. Cancer Biol.* **16**, 253–264 (2006).
21. Dib, A., Gabrea, A., Glebov, O., Bergsagel, P. L. & Kuehl, W. M. Characterization of MYC translocations in multiple myeloma cell lines. *J. Natl Cancer Inst.* (in the press).
22. Liu, J. & Levens, D. Making myc. *Curr. Top. Microbiol. Immunol.* **302**, 1–32 (2006).
23. Garraway, L. A. & Sellers, W. R. Lineage dependency and lineage-survival oncogenes in human cancer. *Nature Rev. Cancer* **6**, 593–602 (2006).
24. Solimini, N. L., Luo, J. & Elledge, S. J. Non-oncogene addiction and the stress phenotype of cancer cells. *Cell* **130**, 986–988 (2007).
25. Hauf, S. *et al.* Dissociation of cohesin from chromosome arms and loss of arm cohesion during early mitosis depends on phosphorylation of SA2. *PLoS Biol.* **3**, e69 (2005).
26. Ge, H. & Roeder, R. G. Purification, cloning, and characterization of a human coactivator, PC4, that mediates transcriptional activation of class II genes. *Cell* **78**, 513–523 (1994).
27. Vassilev, L. T. *et al.* *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**, 844–848 (2004).
28. Polo, J. M. *et al.* Specific peptide interference reveals BCL6 transcriptional and oncogenic mechanisms in B-cell lymphoma cells. *Nature Med.* **10**, 1329–1335 (2004).
29. Alizadeh, A. A. *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This research was supported by the Intramural Research Program of the National Institute of Health (NIH), National Cancer Institute, Center for Cancer Research. Y.Z., B.C. and J.E. were supported by NCI grants CA97513 and CA113992. We wish to thank K. Meyer for her assistance with GEO submissions; D. Levens, J. Liu and H.-J. Chung for assistance with MYC ChIP assay design and the MYC promoter-GFP reporter construct; K. Ozato and L. Ramakrishna for IRF4-deficient mice; and M. Kuehl and members of the Staudt laboratory for their assistance and discussions.

Author Contributions Experimental design/discussion, A.L.S., V.N.N., J.E. and L.M.S.; preparation and performance of experiments, A.L.S., N.C.T.E., L.L., V.N.N., S.D., X.Y., H.Z., Y.Z. and B.C.; data analysis/interpretation, A.L.S., N.C.T.E., W.X., G.W., J.P., J.E. and L.M.S.; manuscript preparation, A.L.S. and L.M.S.

Author Information Microarray data are deposited in the NCBI GEO database under accession numbers GSE8958, GSE9067 (gene expression) and GSE9367 (ChIP-CHIP). Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to L.M.S. (lstaedt@mail.nih.gov).

LETTERS

Essential role for Nix in autophagic maturation of erythroid cells

Hector Sandoval¹, Perumal Thiagarajan², Swapan K. Dasgupta², Armin Schumacher³, Josef T. Prchal⁴, Min Chen¹ & Jin Wang¹

Erythroid cells undergo enucleation and the removal of organelles during terminal differentiation^{1–3}. Although autophagy has been suggested to mediate the elimination of organelles for erythroid maturation^{2–6}, the molecular mechanisms underlying this process remain undefined. Here we report a role for a Bcl-2 family member, Nix (also called Bnip3L)^{7–9}, in the regulation of erythroid maturation through mitochondrial autophagy. *Nix*^{−/−} mice developed anaemia with reduced mature erythrocytes and compensatory expansion of erythroid precursors. Erythrocytes in the peripheral blood of *Nix*^{−/−} mice exhibited mitochondrial retention and reduced lifespan *in vivo*. Although the clearance of ribosomes proceeded normally in the absence of Nix, the entry of mitochondria into autophagosomes for clearance was defective. Deficiency in Nix inhibited the loss of mitochondrial membrane potential ($\Delta\Psi_m$), and treatment with uncoupling chemicals or a BH3 mimetic induced the loss of $\Delta\Psi_m$ and restored the sequestration of mitochondria into autophagosomes in *Nix*^{−/−} erythroid cells. These results suggest that Nix-dependent loss of $\Delta\Psi_m$ is important for targeting the mitochondria into autophagosomes for clearance during erythroid maturation, and interference with this function impairs erythroid maturation and results in anaemia. Our study may also provide insights into molecular mechanisms underlying mitochondrial quality control involving mitochondrial autophagy.

Nix, a BH3-only member of the Bcl-2 family, is upregulated in erythroid cells undergoing terminal differentiation¹⁰. To determine the potential function for Nix in erythroid maturation, we generated *Nix*^{−/−} mice using embryonic stem (ES) cells with a gene trap insertion between exons 3 and 4 of *Nix* (Supplementary Fig. 2). We first examined red blood cells in the peripheral blood (RBCs), including reticulocytes and erythrocytes, in *Nix*^{−/−} mice. Although RBC counts were decreased (Supplementary Table 1), polychromasia and increased reticulocytes were observed in *Nix*^{−/−} mice (Fig. 1a and Supplementary Fig. 3a). We also examined RBCs for the expression of an erythroid cell marker, glycophorin-A-associated Ter119, and for transferrin receptor CD71, which is downregulated during terminal erythroid differentiation^{11,12}. Although Ter119^{low}CD71^{high} and Ter119⁺CD71^{high} early erythroblasts¹³ were absent in the peripheral blood, a significant increase in Ter119⁺CD71⁺ reticulocytes was observed in *Nix*^{−/−} mice (Fig. 1b). Electron microscopy also showed more irregularly shaped cells characteristic of reticulocytes in the peripheral blood of *Nix*^{−/−} mice (Supplementary Fig. 3c). In contrast, CD71[−] mature erythrocytes were reduced in *Nix*^{−/−} mice (Fig. 1b). Together, these data suggest a potential defect in the development from CD71⁺ reticulocytes into mature CD71[−] erythrocytes in *Nix*^{−/−} mice.

Terminal differentiation of reticulocytes into erythrocytes involves the coordinated removal of organelles¹⁴. Because Nix localizes to the

mitochondria¹⁰, we examined whether *Nix*^{−/−} reticulocytes might be defective in eliminating the mitochondria. Consistent with increased reticulocytes, more RBCs (28% at 3 weeks and 12% at 6 weeks of age) were positive for CD71 and mitochondrial staining in *Nix*^{−/−} mice

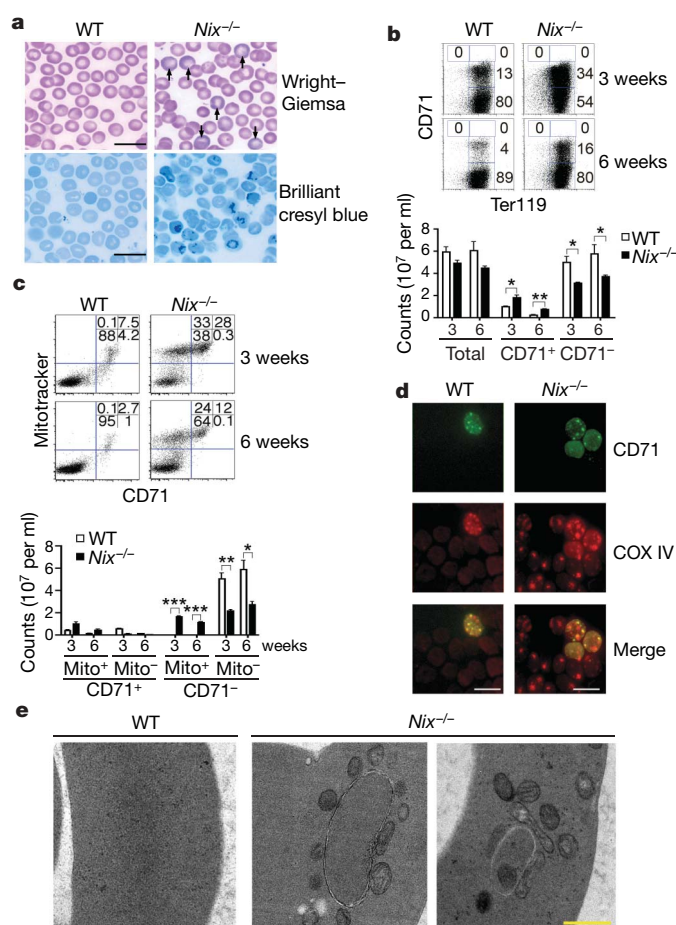


Figure 1 | Reticulocytosis and retention of mitochondria in *Nix*^{−/−} RBCs.

a, Polychromasia (Wright–Giemsa stain, indicated by arrows) and increased reticulocytes (brilliant cresyl blue stain) revealed by staining blood smears from 6-week-old wild-type (WT) or *Nix*^{−/−} mice (scale bar, 20 μ m). **b**, **c**, Ter119 (**b**) or Mitotracker deep red (**c**) versus CD71 staining of RBCs from 3- or 6-week-old mice. Cell counts were also calculated (mean \pm s.e.m.; $n = 3$). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **d**, **e**, Analyses of RBCs as in **b** by immunocytochemistry (**d**; scale bar, 10 μ m) or transmission electron microscopy (**e**; scale bar, 0.5 μ m).

¹Departments of Immunology, ²Pathology and ³Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030, USA. ⁴Division of Hematology, University of Utah School of Medicine, Salt Lake City, Utah 84132, USA.

than in wild-type controls (Fig. 1c). Notably, a significant portion of CD71⁺ RBCs in *Nix*^{-/-} mice retained mitochondria (33% and 24% of total RBCs in 3-week-old and 6-week-old mice, respectively), whereas mitochondria were virtually absent in wild-type CD71⁺ RBCs (Fig. 1c). Immunocytochemistry staining for COX IV, a catalytic subunit of cytochrome *c* oxidase localized to the inner mitochondrial membrane, also confirmed the presence of mitochondria in both CD71⁺ and CD71⁻ RBCs in *Nix*^{-/-} mice (Fig. 1d). Electron microscopy revealed that very few RBCs in wild-type mice contained

mitochondria or detectable autophagosomes (Fig. 1e). Autophagosomal structures were abundant in *Nix*^{-/-} RBCs (Fig. 1e). However, all *Nix*^{-/-} RBCs containing autophagosomes also displayed multiple mitochondria outside of autophagosomes, while approximately 20% of autophagosomes contained one mitochondrion inside (Fig. 1e). This suggests a potential defect in the inclusion of mitochondria by autophagosomes for clearance in *Nix*^{-/-} RBCs.

The loading and unloading of oxygen by haemoglobin can induce oxidant stress in RBCs¹⁵. The mitochondrion is a major site for the production of reactive oxygen species (ROS) and can function as an apoptotic machinery¹⁶. The removal of mitochondria is therefore potentially beneficial for the survival of RBCs. We therefore examined whether *Nix* deficiency could affect RBC survival *in vivo*. Wild-type and *Nix*^{-/-} mice were injected with *N*-hydroxysuccinimido (NHS)-biotin. Cell clearance was determined by the loss of biotinylated RBCs¹⁷. We observed that biotinylated RBCs disappeared more rapidly in *Nix*^{-/-} mice than in wild-type controls (Fig. 2a, top panel), suggesting a faster clearance of RBCs in *Nix*^{-/-} mice. We also labelled RBCs with 5-chloromethylfluorescein diacetate (CMFDA)¹⁸. We then measured the loss of CMFDA-labelled wild-type or *Nix*^{-/-} RBCs after transfer into wild-type recipient mice. We observed accelerated clearance of transferred *Nix*^{-/-} RBCs compared to wild-type controls (Fig. 2a, bottom panel). These data suggest that *Nix*^{-/-} RBCs have a shorter lifespan *in vivo*.

Consistent with decreased survival of *Nix*^{-/-} RBCs *in vivo*, spontaneous caspase activation was observed in *Nix*^{-/-} but not wild-type RBCs after *in vitro* culture (Fig. 2b, c). Moreover, mitochondrion-containing RBCs from *Nix*^{-/-} mice displayed phosphatidylserine, a signal for phagocytosis¹⁹, before and after *in vitro* culture, as revealed by annexin V staining (Fig. 2d). A pan-caspase inhibitor, quinolyl-valyl-O-methylaspartyl-[2,6-difluorophenoxy]-methyl ketone (qVD-opb)²⁰, blocked caspase activation and suppressed annexin V staining on *Nix*^{-/-} RBCs (Fig. 2c, e), indicating that caspase activation contributed to phosphatidylserine display. Increased ROS was detected in *Nix*^{-/-} RBCs, and ROS scavengers inhibited caspase activation in *Nix*^{-/-} RBCs (Fig. 2e). In agreement with accelerated clearance of RBCs, histochemistry showed increased iron-laden macrophages in the spleens of *Nix*^{-/-} mice (Fig. 2f). Decreases in mature erythrocytes have been shown to induce compensatory erythropoiesis¹³. Consistently, reduced mature RBCs were correlated with increased erythropoietin mRNA and an expansion of Ter119⁺CD71⁺ erythroblasts in *Nix*^{-/-} mice (Fig. 2g and Supplementary Fig. 8). The increased erythropoietin may contribute to the expansion of Ter119⁺CD71⁺ erythroblasts, thus ameliorating the anaemia in *Nix*^{-/-} mice. Together, our data suggest that mitochondrial retention in *Nix*^{-/-} RBCs results in increased caspase activation and phosphatidylserine display, leading to faster clearance of RBCs and compensatory erythropoiesis. Removing mitochondria by autophagy would limit mitochondrial disruption in the cytosol and prevent the induction of apoptosis signalling in RBCs.

To induce the generation of sufficient numbers of reticulocytes for analyses, we treated mice with phenylhydrazine (PHZ), which induces reticulocytosis²¹. Ter119⁺CD71⁺ reticulocytes in the peripheral blood were then sorted either on day 3 or on day 6 after PHZ treatment. Notably, a significant portion of the freshly sorted wild-type reticulocytes (18% at day 3 and 37% at day 6 after PHZ treatment) had already lost their mitochondria (Fig. 3a, top row). In contrast, fewer *Nix*^{-/-} reticulocytes were negative for mitochondrial staining (Fig. 3a). This implies that *Nix*^{-/-} reticulocytes were undergoing less efficient elimination of the mitochondria *in vivo*. After culture for *in vitro* maturation³, *Nix*^{-/-} reticulocytes showed significant delays in the removal of the mitochondria compared to wild-type controls (Fig. 3a and Supplementary Fig. 4a). This suggests that *Nix*^{-/-} reticulocytes are defective in clearing mitochondria during maturation.

Inhibition of autophagy with 3-methyladenine, wortmannin or chloroquine^{22,23} suppressed the removal of mitochondria but not

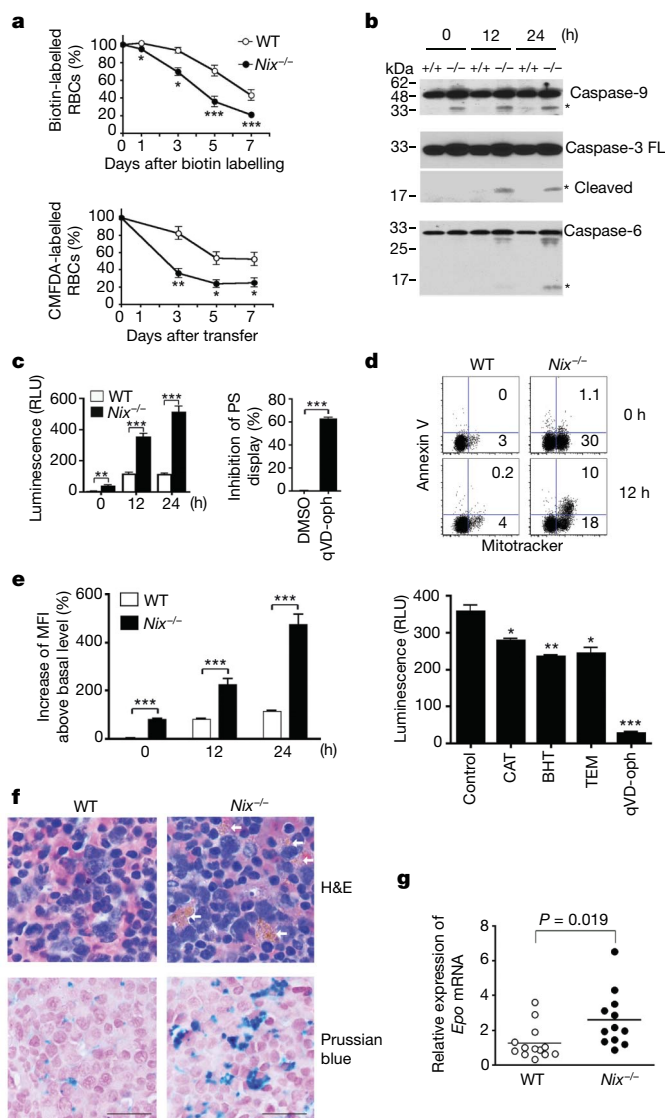


Figure 2 | Decreased survival of RBCs in *Nix*^{-/-} mice. **a**, Quantification of NHS-biotin-labelled RBCs or transferred CMFDA-labelled RBCs ($n = 3$). **b**, Western blot analyses of caspases in RBCs after *in vitro* culture. Asterisks denote processed caspases. FL, full length. **c**, The relative luminescence units (RLU) of caspase activities in cultured RBCs and the suppression of annexin V staining in *Nix*^{-/-} RBCs after 12-h culture in the presence of qVD-opb or solvent control (DMSO) ($n = 3$). **d**, Mitotracker versus annexin V staining of RBCs after *in vitro* culture. **e**, Mean fluorescent intensity (MFI) of ROS staining in RBCs after *in vitro* culture, and caspase activities in *Nix*^{-/-} RBCs after 24 h culture with solvent control, ROS scavengers or qVD-opb ($n = 3$). BHT, butylated hydroxytoluene; CAT, catalase; TEM, tempol. **f**, Haematoxylin and eosin (H&E) staining of spleen sections of 9-week-old wild-type and *Nix*^{-/-} mice. Arrows denote iron deposits within macrophage cytoplasm. Iron deposits were also stained with Prussian blue, followed by counterstain with nuclear-fast red. Scale bar, 20 μ m. **g**, Real-time RT-PCR of erythropoietin (*Epo*) ($n = 12$). For all relevant panels, statistical significance of the data (mean \pm s.e.m.) is: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

ribosomes in reticulocytes (Supplementary Fig. 5), indicating that the clearance of mitochondria, but not ribosomes, is dependent on autophagy in RBCs. Abnormal autophagosome formation in *Nix*^{-/-} reticulocytes might lead to defective mitochondrial clearance. We therefore stained reticulocytes for the processed form of microtubule-associated protein light chain 3 (LC3) that is characteristic of autophagosome formation²⁴. Wild-type and *Nix*^{-/-} reticulocytes had similar levels of LC3 punctate staining at the initiation of culture (Fig. 3c; LC3 dots per cell at day 0: wild type, 19.7 ± 3.9; *Nix*^{-/-}, 21.3 ± 4.5; *P* = 0.19), indicating that the formation of autophagosomes was normal in the absence of Nix. Notably, LC3 co-localized with COX IV in wild-type but not *Nix*^{-/-} reticulocytes (Fig. 3c). Moreover, wild-type but not *Nix*^{-/-} reticulocytes had lost most COX IV staining by day 4 of *in vitro* culture (Fig. 3c). Electron microscopy showed that autophagic vacuoles in wild-type reticulocytes contained engulfed mitochondria; however, most mitochondria remained clustered outside of the autophagic vacuoles in *Nix*^{-/-} reticulocytes (Fig. 3d and Supplementary Fig. 6). This suggests that Nix has a specific role in targeting the mitochondria into autophagosomes for clearance. In contrast to defective mitochondrial removal, *Nix*^{-/-} RBCs showed no retention of the nucleus and underwent the loss of reticular staining similar to wild-type cells (Figs 1a, and 3b and Supplementary Fig. 3c). Therefore, Nix seems to be required for the clearance of mitochondria, but not the nucleus or ribosomes, during erythroid maturation.

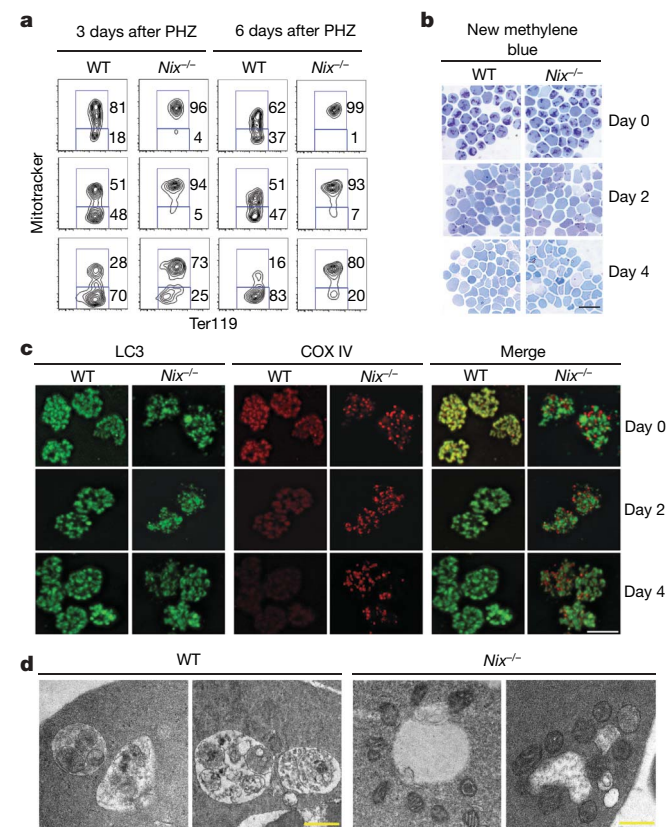


Figure 3 | Defective clearance of mitochondria by autophagy in *Nix*^{-/-} reticulocytes. **a**, Ter119⁺CD71⁺ reticulocytes sorted at day 3 or 6 after PHZ treatment were cultured for *in vitro* maturation for 0 (top row), 2 (middle row) or 4 (bottom row) days, followed by Mitotracker deep red and Ter119 staining. **b**, CD71⁺Ter119⁺ reticulocytes sorted at day 3 after PHZ treatment were cultured and stained with new methylene blue. Scale bar, 20 μ m. **c**, Cells as in **b** were stained for LC3 and COX IV and analysed by deconvolution microscopy. Scale bar, 5 μ m. The Pearson coefficient for LC3 and COX IV co-localization at day 0 (mean \pm s.e.m.) is: wild type, 0.81 \pm 0.001; *Nix*^{-/-}, 0.44 \pm 0.036 (*n* = 35, *P* = 0.0006). **d**, Cells sorted as in **b** were analysed by electron microscopy. Scale bar, 0.5 μ m.

Incubation of Nix with purified mitochondria induced the loss of $\Delta\Psi_m$ (ref. 9). Wild-type reticulocytes showed continued decreases in $\Delta\Psi_m$ that preceded mitochondrial removal during *in vitro* maturation, whereas *Nix*^{-/-} reticulocytes were defective in the loss of $\Delta\Psi_m$ (Supplementary Fig. 7a). This supports a role for Nix in disrupting $\Delta\Psi_m$ during erythroid maturation. Treatment with carbonyl cyanide *p*-trifluoromethoxyphenylhydrazone (FCCP), an uncoupling agent that dissipates the proton gradient across mitochondrial inner membrane to abolish $\Delta\Psi_m$ (ref. 25) (Supplementary Fig. 7c, d), promoted the clearance of mitochondria and restored the entry of the mitochondria into autophagic vacuoles in *Nix*^{-/-} reticulocytes (Fig. 4a, b and Supplementary Fig. 4b). A BH3 mimetic, ABT-737 (ref. 26), which induced the loss of $\Delta\Psi_m$ (Supplementary Fig. 7c), also restored the entry of mitochondria into autophagosomes and promoted mitochondrial removal in *Nix*^{-/-} reticulocytes (Fig. 4a, c and Supplementary Fig. 4b). ABT-737 has been shown to promote the initiation of autophagy by releasing beclin 1 from the suppression mediated by Bcl-2 or Bcl-xL²⁷. Because Nix is not required for autophagosome formation (Fig. 3c), ABT-737 potentially rescued mitochondrial autophagy in *Nix*^{-/-} reticulocytes by inducing the loss of $\Delta\Psi_m$. Although a BH3 mimetic can rescue the defects in *Nix*^{-/-} RBCs, deletion of another BH3-only molecule, Bim, did not affect erythroid differentiation (Supplementary Fig. 8). Therefore, Nix may be a specific sensor triggered by maturation signals in reticulocytes to promote mitochondrial autophagy.

Our study identifies an unexpected function for Nix in mitochondrial autophagy during erythroid differentiation. Nix-dependent loss of $\Delta\Psi_m$ may induce the display of molecules on the outer surface of mitochondria for recognition and sequestration by autophagosomes (Supplementary Fig. 1). Two yeast mitochondrial proteins, Uth1p (ref. 28) and Aup1p (ref. 29), have been shown to regulate mitochondrial autophagy in yeast. Here, we provide an example of specific regulation of mitochondrial autophagy in mammalian cells.

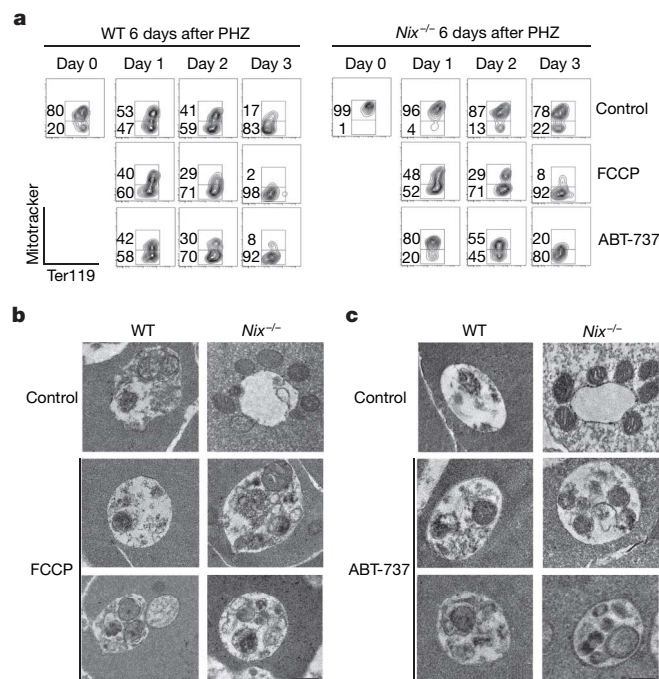


Figure 4 | Removal of mitochondria by autophagy in FCCP-treated or ABT-737-treated *Nix*^{-/-} reticulocytes. **a**, Ter119⁺CD71⁺ reticulocytes were sorted from the peripheral blood of wild-type and *Nix*^{-/-} mice at day 6 after PHZ treatment. The cells were cultured for *in vitro* maturation in the presence of 10 μ M FCCP or 1 μ M ABT-737, followed by staining with Mitotracker deep red and phycoerythrin-conjugated anti-Ter119. **b**, **c**, Reticulocytes were cultured with FCCP (**b**) or ABT-737 (**c**) for 24 h and analysed by transmission electron microscopy. Scale bar, 0.5 μ m.

Defective mitochondrial removal in *Nix*^{-/-} RBCs was associated with spontaneous caspase activation, increased phosphatidylserine display and accelerated clearance by macrophages. Defective autophagy-dependent maturation and increased RBC clearance may contribute to anaemia in *Nix*^{-/-} mice. In human K562 cells that were induced to undergo erythroid maturation, *Nix*-dependent removal of mitochondria was also observed (Supplementary Fig. 9). Defects in autophagic maturation of erythroid cells have been linked to anaemia in both humans and mice^{4,14}. Understanding the mechanisms for mitochondrial autophagy in erythroid cells should facilitate the development of novel therapeutic approaches for treating haematological disorders involving defective erythroid maturation. It may also shed light on the mechanisms underlying mitochondrial quality control by autophagy in the protection against ageing, cancer and neurodegenerative diseases³⁰.

METHODS SUMMARY

Mice. A mouse embryonic stem cell clone with a gene trap insertion targeting the *Nix* locus (The Sanger Institute Gene Trap Consortium) was used to generate *Nix*^{-/-} mice (Supplementary Fig. 2). Mice were backcrossed to the C57BL/6 background for more than five generations.

Analyses of autophagosomes. RBCs were incubated with mouse anti-COX IV (Invitrogen) and a rabbit antibody to processed LC3 (Abgent), followed with Alexa Fluoro-conjugated secondary antibodies (Molecular Probes). The slides were examined by deconvolution microscopy (Applied Precision). RBCs were also fixed and embedded in Spurr's low-viscosity resin, and the sections were prepared and stained with uranyl acetate and lead citrate for analyses by transmission electron microscopy.

Analyses of apoptosis signalling. Cell lysates were used for SDS-PAGE and western blot analysis by probing with antibodies to different caspases, followed by incubation with horseradish peroxidase-conjugated secondary antibodies and developed by the chemiluminescent method (Pierce). Caspase activities were also measured using Ac-DEVD-pNA caspase-Glo reagent (Promega). To measure the display of phosphatidylserine on the cell surface, RBCs labelled with Mitotracker deep red were incubated with FITC-annexin V and analysed by flow cytometry.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 2 December 2007; accepted 16 April 2008.

Published online 4 May 2008.

- Yoshida, H. *et al.* Phosphatidylserine-dependent engulfment by macrophages of nuclei from erythroid precursor cells. *Nature* **437**, 754–758 (2005).
- Fader, C. M. & Colombo, M. I. Multivesicular bodies and autophagy in erythrocyte maturation. *Autophagy* **2**, 122–125 (2006).
- Koury, M. J., Koury, S. T., Kopsombut, P. & Bondurant, M. C. *In vitro* maturation of nascent reticulocytes to erythrocytes. *Blood* **105**, 2168–2174 (2005).
- Kent, G., Minick, O. T., Volini, F. I. & Orfei, E. Autophagic vacuoles in human red cells. *Am. J. Pathol.* **48**, 831–857 (1966).
- Heynen, M. J. & Verwilghen, R. L. A quantitative ultrastructural study of normal rat erythroblasts and reticulocytes. *Cell Tissue Res.* **224**, 397–408 (1982).
- Takano-Ohmuro, H., Mukaida, M., Kominami, E. & Morioka, K. Autophagy in embryonic erythroid cells: its role in maturation. *Eur. J. Cell Biol.* **79**, 759–764 (2000).
- Chen, G. *et al.* *Nix* and *Nip3* form a subfamily of pro-apoptotic mitochondrial proteins. *J. Biol. Chem.* **274**, 7–10 (1999).
- Imazu, T. *et al.* *Bcl-2/E1B* 19 kDa-interacting protein 3-like protein (*Bnip3L*) interacts with *bcl-2/Bcl-xL* and induces apoptosis by altering mitochondrial membrane permeability. *Oncogene* **18**, 4523–4529 (1999).
- Diwan, A. *et al.* Unrestrained erythroblast development in *Nix*^{-/-} mice reveals a mechanism for apoptotic modulation of erythropoiesis. *Proc. Natl Acad. Sci. USA* **104**, 6794–6799 (2007).

- Aerbajinai, W., Giattina, M., Lee, Y. T., Raffeld, M. & Miller, J. L. The proapoptotic factor *Nix* is coexpressed with *Bcl-xL* during terminal erythroid differentiation. *Blood* **102**, 712–717 (2003).
- Kina, T. *et al.* The monoclonal antibody TER-119 recognizes a molecule associated with glycophorin A and specifically marks the late stages of murine erythroid lineage. *Br. J. Haematol.* **109**, 280–287 (2000).
- Pan, B. T. & Johnstone, R. M. Fate of the transferrin receptor during maturation of sheep reticulocytes *in vitro*: selective externalization of the receptor. *Cell* **33**, 967–978 (1983).
- Socolovsky, M. *et al.* Ineffective erythropoiesis in *Stat5a*^{-/-5b}^{-/-} mice due to decreased survival of early erythroblasts. *Blood* **98**, 3261–3273 (2001).
- Holm, T. M. *et al.* Failure of red blood cell maturation in mice with defects in the high-density lipoprotein receptor *SR-BI*. *Blood* **99**, 1817–1824 (2002).
- Sivilotti, M. L. Oxidant stress and haemolysis of the human erythrocyte. *Toxicol. Rev.* **23**, 169–188 (2004).
- Raha, S. & Robinson, B. H. Mitochondria, oxygen free radicals, and apoptosis. *Am. J. Med. Genet.* **106**, 62–70 (2001).
- Hoffmann-Fezer, G. *et al.* Biotin labeling as an alternative nonradioactive approach to determination of red cell survival. *Ann. Hematol.* **67**, 81–87 (1993).
- Levin, J. *et al.* Pathophysiology of thrombocytopenia and anemia in mice lacking transcription factor *NF-E2*. *Blood* **94**, 3037–3047 (1999).
- Fadok, V. A., Bratton, D. L., Frasch, S. C., Warner, M. L. & Henson, P. M. The role of phosphatidylserine in recognition of apoptotic cells by phagocytes. *Cell Death Differ.* **5**, 551–562 (1998).
- Caserta, T. M., Smith, A. N., Gultice, A. D., Reedy, M. A. & Brown, T. L. Q-VD-OPH, a broad spectrum caspase inhibitor with potent antiapoptotic properties. *Apoptosis* **8**, 345–352 (2003).
- Vannucchi, A. M. *et al.* Accentuated response to phenylhydrazine and erythropoietin in mice genetically impaired for their *GATA-1* expression (*GATA-1*^{low} mice). *Blood* **97**, 3040–3050 (2001).
- Lum, J. J. *et al.* Growth factor regulation of autophagy and cell survival in the absence of apoptosis. *Cell* **120**, 237–248 (2005).
- Blommaert, E. F., Krause, U., Schellens, J. P., Vreeling-Sindelarova, H. & Meijer, A. J. The phosphatidylinositol 3-kinase inhibitors wortmannin and LY294002 inhibit autophagy in isolated rat hepatocytes. *Eur. J. Biochem.* **243**, 240–246 (1997).
- Kabeya, Y. *et al.* LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation. *J. Cell Sci.* **117**, 2805–2812 (2004).
- Gottlieb, E., Vander Heiden, M. G. & Thompson, C. B. *Bcl-x(L)* prevents the initial decrease in mitochondrial membrane potential and subsequent reactive oxygen species production during tumor necrosis factor α -induced apoptosis. *Mol. Cell. Biol.* **20**, 5680–5689 (2000).
- Oltersdorf, T. *et al.* An inhibitor of *Bcl-2* family proteins induces regression of solid tumours. *Nature* **435**, 677–681 (2005).
- Maiuri, M. C. *et al.* Functional and physical interaction between *Bcl-X(L)* and a BH3-like domain in *Beclin-1*. *EMBO J.* **26**, 2527–2539 (2007).
- Kissova, I., Deffieu, M., Manon, S. & Camougrand, N. *Uth1p* is involved in the autophagic degradation of mitochondria. *J. Biol. Chem.* **279**, 39068–39074 (2004).
- Tal, R., Winter, G., Ecker, N., Klionsky, D. J. & Abeliovich, H. *Aup1p*, a yeast mitochondrial protein phosphatase homolog, is required for efficient stationary phase mitophagy and cell survival. *J. Biol. Chem.* **282**, 5617–5624 (2007).
- Tatsuta, T. & Langer, T. Quality control of mitochondria: protection against neurodegeneration and ageing. *EMBO J.* **27**, 306–314 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Huang, D. Yoon, A. Syed and D. Townley for technical assistance, and M. Andreeff for ABT-737. This work was supported by grants from the American Society of Hematology (M.C.), the American Heart Association (M.C.) and the NIH (J.W. and J.T.P.), a VA Merit grant (P.T.) and by a Ruth L. Kirschstein National Research Service Award (H.S.).

Author Contributions H.S. conducted the majority of the experiments, supervised by J.W. and M.C.; P.T. stained spleen sections and blood smears; S.K.D. measured osmotic fragility and assisted with biotin and CMFDA labelling; A.S. performed RT-PCR for *Epo*; J.T.P. and P.T. provided experimental advice; M.C. and J.W. generated the *Nix*^{-/-} mice, designed experiments and prepared the manuscript; and all authors edited the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.C. (minc@bcm.tmc.edu) or J.W. (jinwang@bcm.tmc.edu).

METHODS

Mice. A mouse embryonic stem cell clone with a gene trap insertion targeting the *Nix* locus (The Sanger Institute Gene Trap Consortium) was used to generate *Nix*^{-/-} mice (Supplementary Fig. 2). Mice were backcrossed to the C57BL/6 background for more than five generations.

Flow cytometry and immunocytochemistry. Erythroid cells from peripheral blood, bone marrow or spleen of wild-type and *Nix*^{-/-} mice were stained with FITC-anti-CD71 and phycoerythrin-anti-Ter119 (BD Bioscience). Phosphatidylserine display was measured by staining with FITC-annexin V (BD Biosciences). To stain mitochondria, wild-type and *Nix*^{-/-} RBCs were labelled with 200 nM Mitotracker deep red (Molecular Probes) at 37 °C for 30 min. The cells were then stained with FITC-anti-CD71 and phycoerythrin-anti-Ter119 and analysed by flow cytometry. For immunocytochemistry analyses, RBCs from wild-type and *Nix*^{-/-} mice were stained with FITC-anti-CD71 and applied to slides by cytospin. The cells were fixed, incubated with mouse anti-COX IV (Invitrogen) and Alexa Fluor-conjugated secondary antibody. For mitochondria and LC3 co-staining, wild-type and *Nix*^{-/-} RBCs were incubated with mouse anti-COX IV (Invitrogen) and a rabbit antibody to processed LC3 (Abgent), followed by staining with Alexa Fluor-conjugated secondary antibodies (Molecular Probes). The slides were examined using an AxioPlan2 fluorescent microscope (Zeiss) or a SoftWorx Image deconvolution microscope (Applied Precision). The Pearson coefficient of correlation for LC3 and COX IV co-localization was determined by using the SoftWorx software (Applied Precision), with a value greater than 0.5 considered to have co-localization between the two signals.

Measurement of clearance of RBCs. Wild-type and *Nix*^{-/-} mice were injected intravenously with (150 mg kg⁻¹ body weight) NHS-biotin (Pierce), which covalently binds to free amino groups of cell surface proteins¹⁷. Cell clearance was determined by the loss of biotinylated RBCs¹⁷. RBCs were collected at indicated time points and the cells were incubated with phycoerythrin-Streptavidin (BD Bioscience). The percentage of labelled cells was analysed by flow cytometry. RBCs of wild-type or *Nix*^{-/-} mice were also labelled with 4 mM CMFDA (Molecular Probes) which emits green fluorescence after cleavage by intracellular esterases¹⁸. The labelled cells were injected into wild-type recipient mice intravenously. Blood was collected at indicated time points to quantify CMFDA-labelled cells by flow cytometry.

Induction of reticulocytosis in mice. Mice were injected intraperitoneally on day -1 and day 0 with phenylhydrazine (PHZ, Sigma; 60 mg kg⁻¹ body weight).

Ter119⁺CD71⁺ reticulocytes were then sorted on day 3 or day 6 after PHZ treatment using a FACSaria cell sorter (BD Bioscience).

Transmission electron microscopy. RBCs of wild-type and *Nix*^{-/-} mice were fixed and embedded in Spurr's low viscosity resin similar to described protocols³¹. Sections were prepared and stained with uranyl acetate and lead citrate, followed by analyses using a Hitachi H-7500 transmission electron microscope.

Quantitative real-time RT-PCR. RNA was isolated from kidneys of 6-week-old wild-type and *Nix*^{-/-} mice using the MELT total nucleic acid isolation system (Ambion). Erythropoietin mRNA was quantified by real-time RT-PCR and normalized against 18S RNA as described³².

Analyses of apoptosis signalling. Cell lysates were used for SDS-PAGE and western blot analysis by probing with antibodies to Nix (Kamiya), caspase-6, caspase-9 (Cell Signaling), or actin (Santa Cruz Biotechnology), followed by incubation with horseradish peroxidase-conjugated secondary antibodies (Southern Biotechnology). The blots were developed by the chemiluminescent method (Pierce). To detect caspase activities, cells were added to 96-well plates with caspase-Glo reagent containing Ac-DEVD-pNA (Promega) and incubated at room temperature for 2 h. The relative luminescence unit (RLU) was measured in a luminometer (Labsystems). In some experiments, *Nix*^{-/-} RBCs were cultured with solvent control, 4,000 U ml⁻¹ catalase, 100 µM butylated hydroxytoluene, 5 µM tempol or 200 µM qVD-oph for 24 h. Caspase activities were measured as above. To measure the display of phosphatidylserine on the cell surface, RBCs with or without Mitotracker deep red labelling were incubated with FITC-annexin V (BD Biosciences) and analysed by flow cytometry. To measure ROS, wild-type and *Nix*^{-/-} RBCs were cultured *in vitro* and stained with 20 µM 2',7'-dichlorofluorescein diacetate. Mean fluorescent intensity (MFI) was plotted.

Statistical analysis. Data were presented as the mean ± s.e.m. and *P* values were determined by two-tailed Student's *t*-test using GraphPad Prism software version 4.0 for Macintosh.

31. Koury, S. T., Koury, M. J. & Bondurant, M. C. Cytoskeletal distribution and function during the maturation and enucleation of mammalian erythroblasts. *J. Cell Biol.* **109**, 3005–3013 (1989).
32. Mok, H., Mendoza, M., Prchal, J. T., Balogh, P. & Schumacher, A. Dysregulation of ferroportin 1 interferes with spleen organogenesis in polycythaemia mice. *Development* **131**, 4871–4881 (2004).

LETTERS

Imaging the biogenesis of individual HIV-1 virions in live cells

Nolwenn Jouvenet^{1,2}, Paul D. Bieniasz^{1,2} & Sanford M. Simon³

Observations of individual virions in live cells have led to the characterization of their attachment, entry and intracellular transport¹. However, the assembly of individual virions has never been observed in real time. Insights into this process have come primarily from biochemical analyses of populations of virions or from microscopic studies of fixed infected cells. Thus, some assembly properties, such as kinetics and location, are either unknown or controversial^{2–5}. Here we describe quantitatively the genesis of individual virions in real time, from initiation of assembly to budding and release. We studied fluorescently tagged derivatives of Gag, the major structural component of HIV-1—which is sufficient to drive the assembly of virus-like particles⁶—with the use of fluorescence resonance energy transfer, fluorescence recovery after photobleaching and total-internal-reflection fluorescent microscopy in living cells. Virions appeared individually at the plasma membrane, their assembly rate accelerated as Gag protein accumulated in cells, and typically 5–6 min was required to complete the assembly of a single virion. These approaches allow a previously unobserved view of the genesis of individual virions and the determination of parameters of viral assembly that are inaccessible with conventional techniques.

We monitored HeLa cells coexpressing untagged Gag and Gag fused to green fluorescent protein (GFP). This avoids the reported morphological defect seen when Gag-GFP is expressed alone⁷, and electron microscopic studies confirmed that virus-like particles (VLPs) containing Gag and Gag-GFP were morphologically indistinguishable from VLPs containing Gag only (data not shown). Gag was detected as a diffuse signal at 5–6 h after transfection and discrete puncta that apparently localized at the plasma membrane began to appear at 6–7 h after transfection (Supplementary Fig. 1a)⁴. However, the signal from puncta at the plasma membrane was partly masked by the strong diffuse fluorescence emitted by cytoplasmic Gag-GFP (Supplementary Fig. 1b). This diffuse fluorescence diminished relative to the intensity of the fluorescent puncta when imaged with total-internal-reflection fluorescent microscopy (TIR-FM), whose illumination decays exponentially from the coverslip/medium interface with a space constant of about 70 nm (ref. 8) (Supplementary Fig. 1b). Thus, TIR-FM is well suited to monitor the appearance of VLPs at the plasma membrane.

Cells that had generated few (less than 20) Gag puncta at 5–6 h after transfection were observed over the ensuing 30–60 min (Supplementary Fig. 1c). Typically, 50–150 puncta per cell appeared during this period (Supplementary Movie 1). Their behaviour fell into two discrete classes. One class appeared over several minutes (Fig. 1a, top, Fig. 1b, left, and Supplementary Movie 2) and showed little lateral movement during and after emergence (average lateral velocity $0.007 \mu\text{m s}^{-1}$, $n = 25$, Supplementary Fig. 2a). Almost all (99%) of these slowly appearing puncta remained in the TIR field until the end of the imaging period. A second class of Gag puncta

behaved differently, emerging in the TIR field in 5–15 s, remaining there for an average of 33 s ($n = 25$), displaying lateral motility (average velocity $0.195 \mu\text{m s}^{-1}$, $n = 25$), and, in every case, disappearing

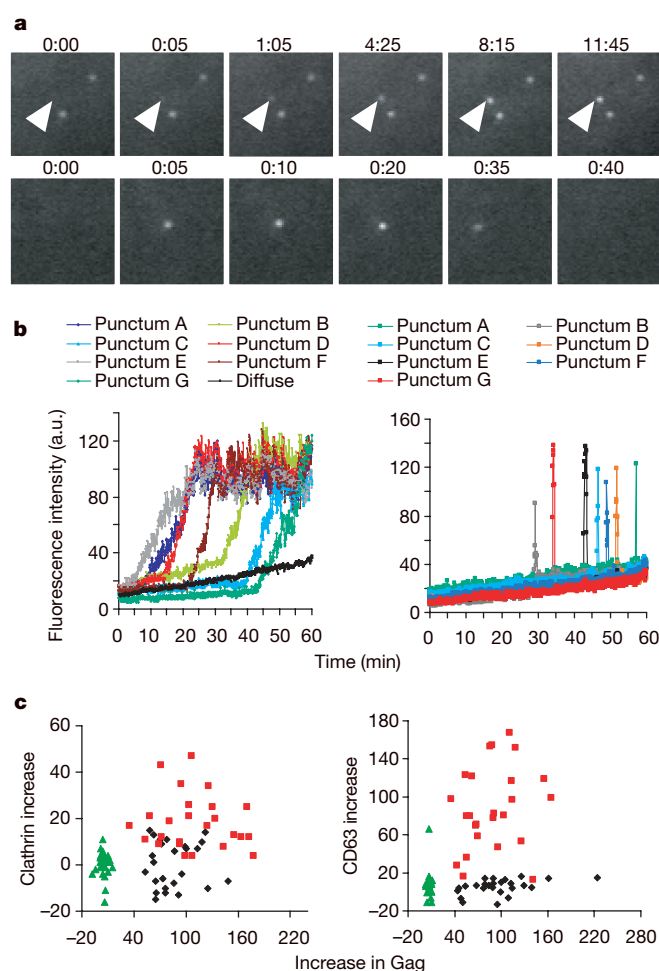


Figure 1 | Two distinct behaviours of HIV-1 Gag puncta at the plasma membrane. **a**, Images illustrating slowly (top panels, arrowheads) and rapidly (bottom panels) appearing Gag-GFP puncta. Fields are $5.5 \mu\text{m} \times 5.5 \mu\text{m}$. Numbers above the fields are minutes:seconds. **b**, Fluorescence intensity in arbitrary units (a.u.) plotted over time for seven slowly appearing puncta (left, coloured lines), for seven rapidly appearing puncta (right) and for seven random areas where no puncta appeared (left, black line). **c**, The change in clathrin (left) or CD63 (right) fluorescence is plotted against the corresponding change in Gag-GFP during the appearance of 25 rapidly emerging (red squares) and slowly emerging (black diamonds) puncta, as well as against the corresponding changes at random areas where no puncta appeared (green triangles).

¹Aaron Diamond AIDS Research Center, ²Laboratory of Retrovirology, and ³Laboratory of Cellular Biophysics, The Rockefeller University, New York, New York 10065, USA.

from the field as rapidly as they appeared (Fig. 1a, bottom, Fig. 1b, right, and Supplementary Movie 3). When the Gag–GFP puncta first appeared (5–6 h after transfection), slowly appearing puncta represented the majority of events (74%, $n = 306$), whereas at later time points (14–15 h after transfection), the rapidly appearing/disappearing population predominated (75%, $n = 166$) (Supplementary Fig. 2b). When Gag was co-expressed with the viral protein Vpu, which reduces the surface tethering of virions and their subsequent appearance in late endosomes^{9–11}, the fraction of rapidly appearing/disappearing puncta fell both at early (from 26% to 3%) and late (from 75% to 34%; $P < 0.001$) time points (Supplementary Fig. 2b).

Virions can appear in late endosomes as a result of active antiviral tethering^{9–11} and subsequent internalization^{4,9–12}. Further, endosomes marked with fluorescently tagged versions of resident proteins, such as CD63 (ref. 13), or coat proteins, such as clathrin^{14,15}, have previously been observed moving in the TIR field. To examine whether the Gag puncta were associated with endosomes, Gag–GFP together with CD63–mCherry or with DsRed–clathrin light chain were detected with simultaneous dual-colour TIR-FM (Supplementary Movie 4). Most of the rapidly appearing/disappearing, laterally motile Gag puncta were positive for CD63 (88%, $n = 25$) and clathrin (72%, $n = 25$). The appearance, movement and disappearance of these puncta were indistinguishable from the CD63-positive or clathrin-positive vesicles (Supplementary Fig. 3a and Supplementary Movies 5 and 6). On average, the fluorescence of CD63 and clathrin increased 16-fold and 50-fold, respectively, coincident with the rapidly appearing Gag puncta, in comparison with randomly chosen, equivalently sized areas of observation containing only diffuse Gag ($P < 0.001$) (Fig. 1c). In contrast, none of the slowly appearing puncta were associated with a detectable increase in CD63

($n = 25$) or clathrin ($n = 25$), and in this respect they were indistinguishable from areas of the plasma membrane containing only diffuse Gag ($P = 0.43$ and 0.98 , respectively) (Supplementary Fig. 3b and Supplementary Movies 7 and 8). Because the rapidly appearing/disappearing puncta were indistinguishable from endosomes, we focused on the slowly appearing puncta as the population that might represent genuine VLP assembly events.

During virion assembly, Gag molecules become sufficiently close to one another that FRET can occur. Indeed, fluorescence resonance energy transfer (FRET) has been demonstrated between retroviral Gag–cyan fluorescent protein (CFP) and Gag–yellow fluorescent protein (YFP) molecules in whole cells^{16–18}. To monitor the assembly of individual particles, we chose the comparatively weak FRET pair of GFP and mCherry¹⁹, precisely because their absorption and emission spectra overlap less than those of CFP and YFP. FRET should therefore be more dependent on the high density of about 5,000 Gag molecules²⁰ found within a single virion. FRET was detected between Gag–GFP and Gag–mCherry, both within Gag-expressing cells and within individual cell-free VLPs (the FRET is characterized in Supplementary Fig. 4). When single Gag puncta first became visible, the FRET coefficient (the ratio of mCherry to GFP) within them was similar to the FRET coefficient measured in areas containing only diffuse Gag (Fig. 2a). Then, both the GFP and mCherry fluorescence increased, with mCherry fluorescence increasing relatively more rapidly until they reached their maxima synchronously (Fig. 2b, c, and Supplementary Fig. 5). Thus, the FRET in appearing Gag puncta increased 2.5-fold to the level in cell-free VLPs (Fig. 2a). Bleaching the mCherry increased the GFP emission, confirming that FRET occurred in VLPs (Supplementary Fig. 4d). Thus, during the slow emergence, Gag molecules achieve greater proximity, the expected signature of an assembly event, and ultimately are as closely packed as in VLPs released into the extracellular milieu.

The stable fluorescence ultimately achieved by each slowly appearing punctum (Fig. 1b, left) could reflect a steady state in which Gag–GFP is exchanging with the cytoplasmic pool, or it may represent the completion of the assembly of a VLP whose Gag molecules are segregated from the cytoplasmic pool. To distinguish between these possibilities, the puncta were examined with fluorescence recovery after photobleaching (FRAP). This analysis revealed that puncta with a low but rising intensity during the pre-bleach period recovered well during the post-bleach period. Conversely, puncta whose intensity was high and stable during the pre-bleach period did not recover (Fig. 3a, b, and Supplementary Fig. 6). Thus, most puncta recruiting Gag molecules before the bleach continued to recruit new Gag molecules after the bleach; those with steady fluorescence represented puncta in which Gag recruitment was completed and irreversible.

The final step in the genesis of an HIV-1 particle is the fission of virion and cell membranes²¹. At this point, the virion interior should not be able to exchange any molecules—even protons—with the cell cytoplasm. To test whether this final step occurred in the Gag puncta we were observing, we expressed Gag fused to pHluorin²², a variant of GFP not fluorescent at acidic pH, and then varied the pCO_2 in the medium²³. Raising the pCO_2 produces acidification as a result of the reaction $CO_2 + H_2O \leftrightarrow H^+ + HCO_3^-$. This should acidify the cytosol more rapidly than the interior of cell-free VLPs as result of the cytosolic enzyme carbonic anhydrase²³. Thus, increasing the pCO_2 should quench the fluorescence of cytosol-exposed Gag–pHluorin more than Gag–pHluorin in a VLP whose lipid envelope has separated from the cell. As expected, after a brief pulse of higher pCO_2 the fluorescence of the cytosolic-exposed diffuse Gag–pHluorin was quenched more than the Gag–pHluorin in cell-free VLPs (Fig. 3c–f). After the pCO_2 pulse, the fluorescence of the diffuse Gag–pHluorin returned to normal more quickly than that of the isolated VLPs (Fig. 3d). The fluorescence of one population of cell-associated puncta responded indistinguishably from the diffuse Gag to a pulse of pCO_2 (Fig. 3e, f). Conversely, a minority of Gag puncta showed the lower sensitivity to pCO_2 of the budded, cell-free VLPs.

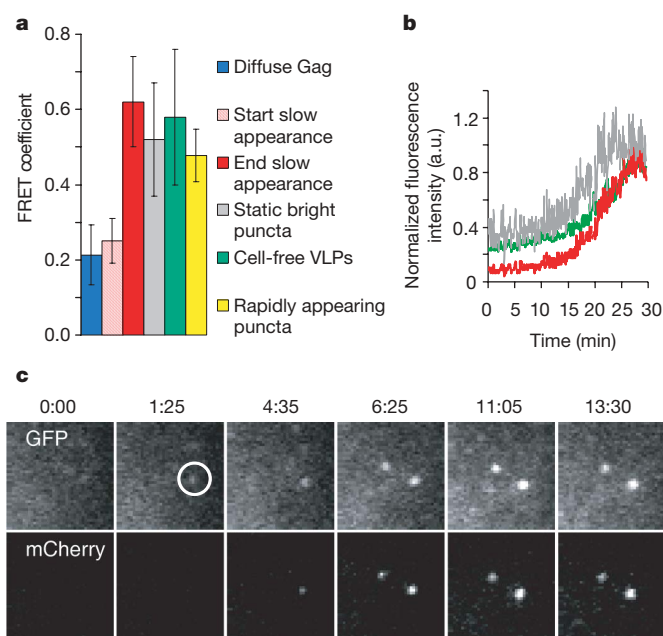


Figure 2 | FRET analysis of individual Gag puncta. Untagged Gag, Gag–mCherry and Gag–GFP were co-expressed and the cells were illuminated with a 488-nm laser. **a**, FRET coefficients measured at the beginning and the end of the slow appearance of individual puncta ($n = 30$), compared with FRET coefficients at randomly chosen areas containing diffuse Gag ($n = 30$), in static bright puncta ($n = 30$), in cell-free VLPs ($n = 100$) or in rapidly appearing puncta ($n = 30$). Errors are s.d. **b**, Plots of the normalized fluorescence intensity of GFP (green), corrected mCherry (red) and the FRET coefficient (grey) during the appearance of the punctum shown in **c**. Additional examples are shown in Supplementary Fig. 5. **c**, Images of GFP and corrected mCherry fluorescence (due to FRET) during the slow appearance of a punctum. Fields are $5.5 \mu m \times 5.5 \mu m$. Numbers above the fields are minutes:seconds.

However, 30 min later, most of the cell-associated Gag puncta had shifted to the less pCO₂-sensitive state of the budded VLPs. (Fig. 3e). In addition, after the pulse of pCO₂, the fluorescence of the diffuse Gag and the more pCO₂-responsive population of Gag puncta returned to normal more quickly than the less responsive puncta or the cell-free VLPs (Fig. 3d). Thus, the population of Gag puncta that were as sensitive as diffuse Gag to cytosolic acidification probably consisted of nascent VLPs whose interiors were continuous with the cytosol, whereas the population relatively resistant to pCO₂ probably consisted of budded VLPs. Moreover, when a Gag-pHluorin protein lacking a functional late-budding domain²⁴ was used to generate puncta, only the pCO₂-sensitive population was seen (Fig. 3f), showing that fission of the VLP and cell membranes was needed to generate pCO₂-resistant (budded) VLPs.

Despite the conclusion that these VLPs have budded and are separate from the cell, they do not move away. This is a confluence of at least two factors. The VLPs are about 100 nm in diameter and their movement is limited by the distance from the cell to the coverslip (20–40 nm). Second, a large fraction of virions remain associated with the cell surface after assembly¹¹.

We have observed a population of laterally non-motile VLPs that appear over several minutes at the plasma membrane, whose emergence is accompanied by a recruitment of Gag molecules that become progressively closer to each other until they segregate from the cytoplasmic pool and detach from the plasma membrane. These slowly appearing VLPs are not associated with endosomes that approach

and retreat from the plasma membrane and represent the vast majority of those that appear at the plasma membrane within the first few hours after Gag expression, particularly in the presence of Vpu. We conclude that this population of slowly appearing puncta represents the genesis of VLPs through *de novo* assembly. Analysis of 370 such assembly events revealed that individual VLPs assembled over an average of 8.5 min, with 5–6 min per VLP being the most frequent rate of assembly (Fig. 4a). Assembly was slower for VLPs that appeared early in the observation period, but accelerated thereafter as Gag concentration increased, stabilizing at 5–6 min (Fig. 4b). Gag concentration influences both the interaction between Gag molecules and the interaction between Gag and the plasma membrane²⁵ and is very probably a key determinant of assembly kinetics. It is possible that some VLP assembly also takes place on internal membranes that are out of the field of view afforded by TIR-FM, although some previous work argues strongly against this notion^{4,5,12}. In examining tens of thousands of VLPs assembling in hundreds of cells, we never observed an organelle reaching the surface and discharging multiple VLPs. Thus, if there is any assembly on internal organelles it does not contribute significantly to the population of VLPs appearing at the basal surface of the cell during the first day of expression or infection.

Altering the expression of Gag-GFP relative to Gag, between ratios of 1:1 and 1:10, did not affect the kinetics of VLP assembly, indicating that their genesis was not influenced by the presence of a GFP tag (Supplementary Fig. 6). Moreover, when Gag was fused to mCherry, which matures more slowly than GFP^{26,27}, VLPs assembled with

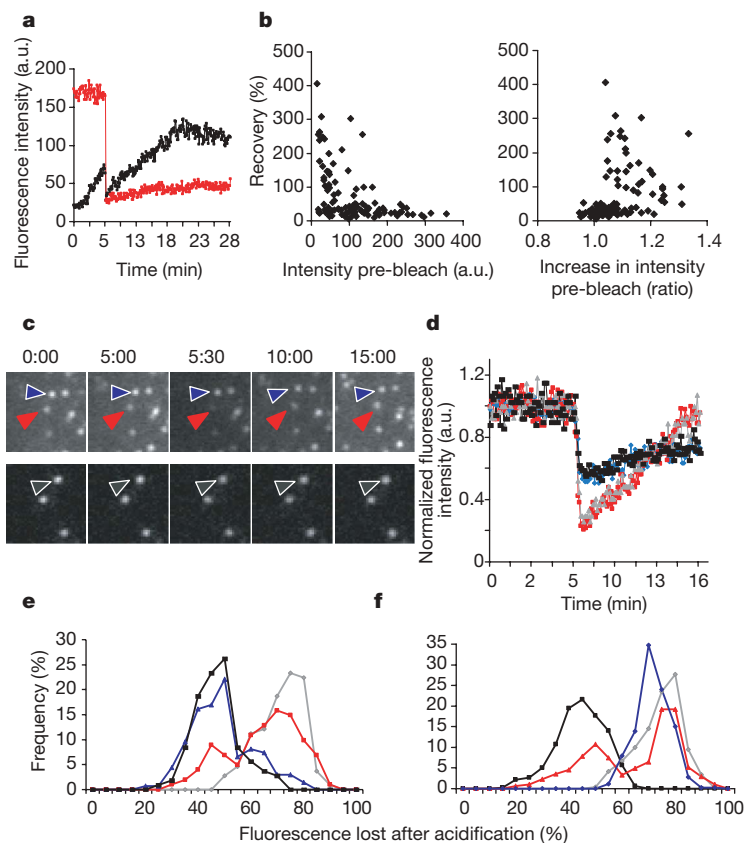


Figure 3 | FRAP analysis of individual Gag puncta. **a**, Plots of Gag-GFP fluorescence intensity over time for the black and red puncta (black and red lines, respectively) shown in Supplementary Fig. 6. **b**, Fluorescence recovery of 100 puncta plotted as a function of either their absolute pre-bleach intensity (left panel) or the fold increase in their intensity before bleaching (right panel) (1 means no increase, more than 1 means increasing). **c**, Images showing Gag-pHluorin puncta before and after the CO₂-mediated acidification. Cells (top panels) or cell-free VLPs (bottom panels) were incubated with pCO₂ for 30 s after 5 min of imaging. The arrows indicate one pCO₂-sensitive punctum (red), one pCO₂-insensitive punctum (blue) and

one cell-free VLP (black). Fields are 3.5 $\mu\text{m} \times 3.5 \mu\text{m}$. Numbers above the fields are minutes:seconds. **d**, Plots of normalized fluorescence intensity over time for the three Gag-pHluorin puncta in **c** (bearing the same colour coding), as well as for a diffuse Gag area (grey). **e**, Loss of fluorescence immediately after pCO₂-mediated acidification for diffuse Gag (grey, $n = 107$), cell-free VLPs (black, $n = 107$) and Gag puncta at $t = 0$ min (red, $n = 101$) and $t = 30$ min (blue, $n = 136$) of imaging. **f**, Loss of fluorescence immediately after pCO₂-mediated acidification of diffuse Gag (grey, $n = 728$), cell-free VLPs (black, $n = 334$), wild-type Gag (red, $n = 308$) and Gag with a late-domain mutation (blue, $n = 366$).

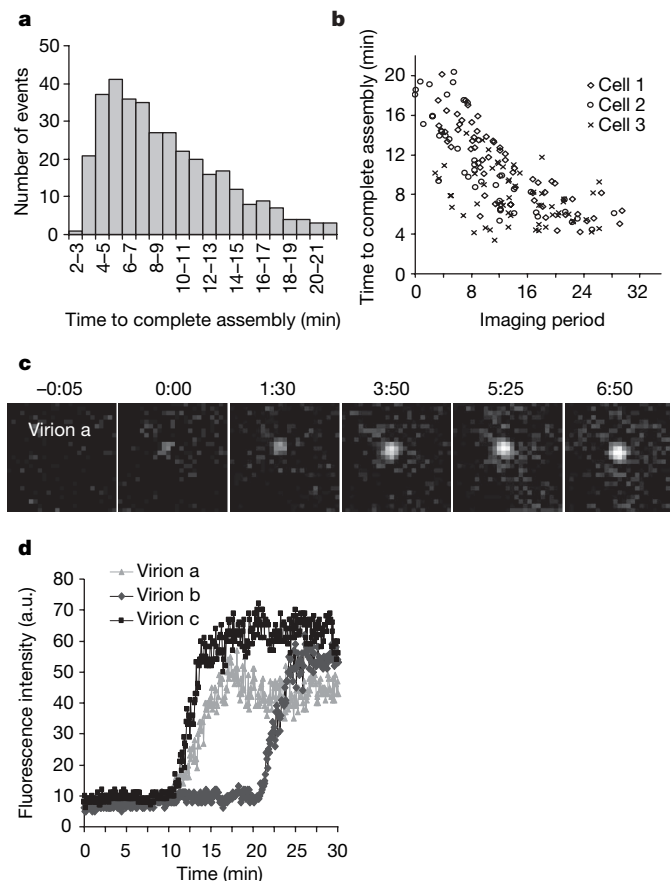


Figure 4 | Variation in HIV-1 assembly kinetics. **a**, Distribution of the time to complete assembly for 370 individual VLPs in 11 HeLa cells expressing Gag/Gag–GFP. Time to complete assembly was defined as the interval between the points of inflection on plots of fluorescence intensity against time, for each VLP (for example Fig. 1b, c). **b**, Time to complete assembly is plotted against the time at which assembly commenced in three cells imaged for 40 min. Zero time is defined as the time at which observation began; that is, when more than 1 but fewer than 20 VLPs were visible in the TIR field for each cell. $R = -0.71$; $R^2 = 0.5$. **c**, **d**, Assembly of HIV-1 virions from full-length proviral plasmids. **c**, Images of an individual HIV-1 virion assembly event. Fields are $5.5 \mu\text{m} \times 5.5 \mu\text{m}$. Numbers above the fields are minutes:seconds. **d**, Plots of the fluorescence intensity over time for three assembly events, including that shown in **c**.

kinetics indistinguishable from those containing Gag/Gag–GFP (Supplementary Fig. 7), indicating that the kinetics was not affected by the maturation time of the fluorophore. Significantly, in cells transfected with full-length proviral plasmids in which YFP was inserted into the stalk region of matrix⁴ (Supplementary Fig. 8), the assembly of HIV-1 virions was completed with similar kinetics to that of Gag–GFP-containing VLPs (Fig. 4c, d, and Supplementary Fig. 7c). Thus, neither the location nor the nature of the fluorophore linked to Gag, nor the presence of other viral proteins, affected the kinetics of particle assembly.

These studies have allowed the observation and quantitative analysis of the assembly of individual virions in living cells. Elaboration of these approaches should permit characterization of the recruitment of other viral and cellular components to nascent virions and could address many questions that are difficult to tackle with conventional techniques.

METHODS SUMMARY

Expression vectors. Plasmids expressing HIV-1 Gag and Gag–GFP, Vpu–YFP and CD63–Cherry proteins were previously described^{4,9}, as was an HIV-1 proviral plasmid that carries YFP embedded within Gag⁴. Derivatives were constructed by using standard molecular biology techniques.

Cells and transfection. HeLa cell clones expressing CD63–mCherry or DsRed–clathrin were derived by retroviral transduction. For imaging, cells plated in glass-bottomed dishes were transfected with Lipofectamine2000. VLPs isolated from the supernatant of transfected cells were adhered to polylysine-coated dishes and imaged under the same conditions as VLP-producing cells.

Image acquisition and analysis. Through-the-objective TIR-FM was performed with an inverted Olympus IX-70 microscope with a 60 \times , 1.45 numerical aperture TIR objective. Imaging of GFP, mCherry or DsRed was achieved by excitation with a 488-nm laser line of an argon laser or a 543-nm HeNe laser, as appropriate. Simultaneous dual-colour imaging was achieved with a dual emission splitter equipped with a 515/30 bandpass filter and a 580lp filter. For FRET analysis, experimentally determined GFP emission bleed-through and direct excitation of mCherry by the 488-nm laser was subtracted from the raw mCherry fluorescence values to yield a corrected mCherry fluorescence that was due to FRET. The FRET coefficient was calculated by dividing the intensity of the corrected mCherry emission by the intensity of the GFP emission in the area of interest. For the analysis of fluorescence recovery, cells were imaged at 5 h after transfection for 5 min, photobleached for 3 min and observed for a further 20 min. During the bleach, the opening of an iris placed in front of the TIR-FM laser was reduced to bleach only a part of the cell. To test the sensitivity of VLPs assembled with Gag–pHluorin to pCO₂, cultures were perfused with 100% CO₂ for 30 s. All image analysis was performed with MetaMorph software.

Received 13 March; accepted 10 April 2008.

Published online 25 May 2008.

1. Brandenburg, B. & Zhuang, X. Virus trafficking—learning from single-virus tracking. *Nature Rev. Microbiol.* **5**, 197–208 (2007).
2. Pelchen-Matthews, A., Kramer, B. & Marsh, M. Infectious HIV-1 assembles in late endosomes in primary macrophages. *J. Cell Biol.* **162**, 443–455 (2003).
3. Sherer, N. M. *et al.* Visualization of retroviral replication in living cells reveals budding into multivesicular bodies. *Traffic* **4**, 785–801 (2003).
4. Jouvénat, N. *et al.* Plasma membrane is the site of productive HIV-1 particle assembly. *PLoS Biol.* **4**, e435 (2006).
5. Welsch, S. *et al.* HIV-1 buds predominantly at the plasma membrane of primary human macrophages. *PLoS Pathogens* **3**, e36 (2007).
6. Gottlinger, H. G. The HIV-1 assembly machine. *AIDS* **15** (Suppl. 5), S13–S20 (2001).
7. Larson, D. R., Johnson, M. C., Webb, W. W. & Vogt, V. M. Visualization of retrovirus budding with correlated light and electron microscopy. *Proc. Natl Acad. Sci. USA* **102**, 15453–15458 (2005).
8. Jaiswal, J. K. & Simon, S. M. Imaging single events at the cell membrane. *Nature Chem. Biol.* **3**, 92–98 (2007).
9. Neil, S. J., Eastman, S. W., Jouvénat, N. & Bieniasz, P. D. HIV-1 Vpu promotes release and prevents endocytosis of nascent retrovirus particles from the plasma membrane. *PLoS Pathogens* **2**, e39 (2006).
10. Neil, S. J., Sandrin, V., Sundquist, W. & Bieniasz, P. D. An interferon- α -induced tethering mechanism inhibits HIV-1 and Ebola virus particle release but is counteracted by the HIV-1 Vpu protein. *Cell Host Microbe* **2**, 193–203 (2007).
11. Neil, S. J., Zang, T. & Bieniasz, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**, 425–430 (2008).
12. Finzi, A., Orthwein, A., Mercier, J. & Cohen, E. A. Productive human immunodeficiency virus type 1 assembly takes place at the plasma membrane. *J. Virol.* **81**, 7476–7490 (2007).
13. Jaiswal, J. K., Andrews, N. W. & Simon, S. M. Membrane proximal lysosomes are the major vesicles responsible for calcium-dependent exocytosis in nonsecretory cells. *J. Cell Biol.* **159**, 625–635 (2002).
14. Keyel, P. A., Watkins, S. C. & Traub, L. M. Endocytic adaptor molecules reveal an endosomal population of clathrin by total internal reflection fluorescence microscopy. *J. Biol. Chem.* **279**, 13190–13204 (2004).
15. Rappoport, J. Z., Taha, B. W. & Simon, S. M. Movement of plasma-membrane-associated clathrin spots along the microtubule cytoskeleton. *Traffic* **4**, 460–467 (2003).
16. Derdowski, A., Ding, L. & Spearman, P. A novel fluorescence resonance energy transfer assay demonstrates that the human immunodeficiency virus type 1 Pr55Gag 1 domain mediates Gag–Gag interactions. *J. Virol.* **78**, 1230–1242 (2004).
17. Hubner, W. *et al.* Sequence of human immunodeficiency virus type 1 (HIV-1) Gag localization and oligomerization monitored with live confocal imaging of a replication-competent, fluorescently tagged HIV-1. *J. Virol.* **81**, 12596–12607 (2007).
18. Larson, D. R., Ma, Y. M., Vogt, V. M. & Webb, W. W. Direct measurement of Gag–Gag interaction during retrovirus assembly with FRET and fluorescence correlation spectroscopy. *J. Cell Biol.* **162**, 1233–1244 (2003).
19. Tramier, M., Zahid, M., Mevel, J. C., Masse, M. J. & Coppey-Moisán, M. Sensitivity of CFP/YFP and GFP/mCherry pairs to donor photobleaching on FRET determination by fluorescence lifetime imaging microscopy in living cells. *Microsc. Res. Tech.* **69**, 933–939 (2006).
20. Briggs, J. A. *et al.* The stoichiometry of Gag protein in HIV-1. *Nature Struct. Mol. Biol.* **11**, 672–675 (2004).

21. Pornillos, O., Garrus, J. E. & Sundquist, W. I. Mechanisms of enveloped RNA virus budding. *Trends Cell Biol.* **12**, 569–579 (2002).
 22. Miesenböck, G., De Angelis, D. A. & Rothman, J. E. Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* **394**, 192–195 (1998).
 23. Simon, S., Roy, D. & Schindler, M. Intracellular pH and the control of multidrug resistance. *Proc. Natl Acad. Sci. USA* **91**, 1128–1132 (1994).
 24. Bieniasz, P. D. Late budding domains and host proteins in enveloped virus release. *Virology* **344**, 55–63 (2006).
 25. Perez-Caballero, D., Hatzioannou, T., Martin-Serrano, J. & Bieniasz, P. D. Human immunodeficiency virus type 1 matrix inhibits and confers cooperativity on gag precursor-membrane interactions. *J. Virol.* **78**, 9560–9563 (2004).
 26. Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnol.* **24**, 79–88 (2006).
 27. Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. A guide to choosing fluorescent proteins. *Nature Methods* **2**, 905–909 (2005).
- Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.
- Acknowledgements** We thank T. Zang for sharing the HeLa cell line stably expressing DsRed-clathrin light chain; A. Baraff for statistical analysis; members of the Bieniasz and Simon laboratories for discussions; and R. Y. Tsien, T. Kirchhausen and G. Miesenböck for plasmids. This work was supported by grants from the National Institutes of Health (to P.D.B. and S.M.S.) and the National Science Foundation (to S.M.S.). N.J. is supported by an amfAR Mathilde Krim Fellowship in Basic Biomedical Research.
- Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.D.B. (pbienias@adarc.org) or S.M.S. (simon@rockefeller.edu).

Mei-P26 regulates microRNAs and cell growth in the *Drosophila* ovarian stem cell lineage

Ralph A. Neumüller¹, Joerg Betschinger^{1†}, Anja Fischer¹, Natascha Bushati², Ingrid Poernbacher¹, Karl Mechtler^{1,3}, Stephen M. Cohen^{2†} & Juergen A. Knoblich¹

Drosophila neuroblasts¹ and ovarian stem cells^{2,3} are well characterized models for stem cell biology. In both cell types, one daughter cell self-renews continuously while the other undergoes a limited number of divisions, stops to proliferate mitotically and differentiates. Whereas neuroblasts segregate the Trim-NHL (tripartite motif and Ncl-1, HT2A and Lin-41 domain)-containing protein Brain tumour (Brat) into one of the two daughter cells^{4–6}, ovarian stem cells are regulated by an extracellular signal from the surrounding stem cell niche. After division, one daughter cell loses niche contact. It undergoes 4 transit-amplifying divisions to form a cyst of 16 interconnected cells that reduce their rate of growth and stop to proliferate mitotically. Here we show that the Trim-NHL protein Mei-P26 (refs 7, 8) restricts growth and proliferation in the ovarian stem cell lineage. Mei-P26 expression is low in stem cells but is strongly induced in 16-cell cysts. In *mei-P26* mutants, transit-amplifying cells are larger and proliferate indefinitely leading to the formation of an ovarian tumour. Like *brat*, *mei-P26* regulates nucleolar size and can induce differentiation in *Drosophila* neuroblasts, suggesting that these genes act through the same pathway. We identify Argonaute-1, a component of the RISC complex, as a common binding partner of Brat and Mei-P26, and show that Mei-P26 acts by inhibiting the microRNA pathway. Mei-P26 and Brat have a similar domain composition that is also found in other tumour suppressors and might be a defining property of a new family of microRNA regulators that act specifically in stem cell lineages.

When *Drosophila* germline stem cells divide (Fig. 1a), 1 daughter cell (the cystoblast) loses niche contact and undergoes 4 transit-amplifying divisions to create a cyst of 16 cystocytes, which remain connected by the fusome. In each cyst, 1 cell becomes the oocyte, whereas the other cells undergo endoreplication to form 15 so-called nurse cells.

Because *brat*¹⁹² mutant germline clones do not show any obvious defects in oogenesis (data not shown), we analysed the ovarian phenotype of *mei-P26*. Like Brat, Mei-P26 is a Trim-NHL protein⁹ and carries an NHL domain, a coiled-coil region and several B-boxes. Weak *mei-P26* mutants have defects in meiosis⁷ whereas stronger alleles cause tumorous overproliferation⁸. *mei-P26* mutant ovaries were stained using the monoclonal antibody mAb1B1 (ref. 10), which labels the fusome and the spectrosome—a cytoplasmic organelle only present in stem cells and cystoblasts¹¹ (Fig. 1b–f). Wild-type ovaries contain two stem cells near the tip of the germarium whereas cystoblasts and cysts are separated from the stem cell niche (Fig. 1b). In *mei-P26* mutant ovaries, germaria are filled with individual spectrosome-containing cells (19%) and cysts containing varying numbers of cells (<4 in 39%, >4 in 42%; *n* > 400 cells) connected by a fusome (Fig. 1c, d and Supplementary Fig. 1a, a'). Nurse cells and

oocytes are not formed, and fusomes or spectrosomes are maintained at later stages of oogenesis (Fig. 1e, f). In wild-type ovaries, the S-phase marker Cyclin E oscillates with the cell cycle in stem cells and mitotically active cysts, is downregulated as cystocytes exit mitotic proliferation and is re-expressed as nurse cells enter endoreplication¹² (Fig. 1g). In *mei-P26* mutant ovaries, however, Cyclin E is highly expressed at all stages of oogenesis (Fig. 1h). Upregulation of Cyclin E occurs even in small *mei-P26* clones (Supplementary Fig. 1b, c), and is therefore not an indirect consequence of tumour formation. Phospho-histone-H3-positive mitotic germline cells are restricted to the anterior tip of the germarium in wild-type ovaries (Fig. 1i) but are detected throughout the ovarioles in *mei-P26* mutants (Fig. 1j and Supplementary Fig. 1a, a'). Whereas all cells in a wild-type cyst divide synchronously, *mei-P26* mutant cysts frequently contain both mitotic and interphase cells (data not shown). Thus, *mei-P26* is required for proliferation control and differentiation in the female germline stem cell lineage.

To test stem cell niche signalling, we used a green fluorescent protein (GFP) fusion to the *bag of marbles* (*bam*) promoter¹³, which is suppressed by Decapentaplegic (Dpp) from the niche in stem cells but not in cystocytes (Fig. 1k and Supplementary Fig. 1d). In *mei-P26* mutants (Fig. 1l and Supplementary Fig. 1e), *bam* transcription is repressed in niche-contacting cells but is upregulated in cystocytes. Because staining for the niche marker Armadillo (Fig. 1m, n) reveals no structural abnormalities and because ovarian tumours are also observed when *mei-P26* is removed exclusively from the female germ line (Supplementary Fig. 1f, g), *mei-P26* is not required for sending or receiving the niche signal.

mei-P26 mutant tumour cells have branched fusomes and express Bam. To confirm further that these cells do not have stem cell identity, we analysed the expression of ooligo RNA-binding protein (Orb). Orb expression starts in all cystocytes between the 8- and 16-cell stage¹⁴, but is restricted to the oocyte during later stages (Fig. 1o). In *mei-P26* mutants, Orb expression is initiated normally (Fig. 1p) and is detected throughout the tumour. However, Orb is never restricted to a single cell, suggesting that the oocyte is not specified. Thus, *mei-P26* mutants develop a 'cystocytoma' in which tumour cells express markers for the cystocyte fate. Single-spectrosome-containing cells in the tumour might arise from occasional disintegration of fusome-connected cysts, a process that has been described before¹⁵.

Brat can inhibit cell growth and ribosome biogenesis¹⁶. *mei-P26* might also regulate cell growth because overexpression from the *eyeless* promoter reduces eye size (Supplementary Fig. 2a, b), even when cell death is inhibited by co-expressing *p35* (data not shown). To test whether cell growth is differentially regulated in the ovarian stem cell lineage, we quantified the volume of stem cells and cysts

¹Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Dr Bohr Gasse 3, 1030 Vienna, Austria. ²European Molecular Biology Laboratory (EMBL), Meyerhofstraße 1, 69117 Heidelberg, Germany. ³Research Institute of Molecular Pathology (IMP), Dr Bohr-Gasse 7, 1030 Vienna, Austria. †Present addresses: Wellcome Trust Centre for Stem Cell Research (CSCR), University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK (J.B.); Temasek Life Sciences Laboratory, 1 Research Link, National University of Singapore, 117604, Singapore (S.M.C.).

after three-dimensional reconstruction (see Methods). Shortly after stem cell division (elongated spectrosome morphology), stem cells are $314 \pm 13 \mu\text{m}^3$ (mean \pm s.e.m.; $n = 4$) whereas cystoblasts are $330 \pm 30 \mu\text{m}^3$ ($n = 4$). Stem cells grow to a maximum of $600 \mu\text{m}^3$ (average $437 \pm 21 \mu\text{m}^3$, $n = 19$) and double their volume between each mitotic division (about once per day). To measure cyst volumes, mitotic clones were marked by the absence of GFP (Fig. 2a). Three- and four-day-old 16-cell cysts are $1,215 \pm 61 \mu\text{m}^3$ ($n = 7$) and $1,163 \pm 48 \mu\text{m}^3$ ($n = 7$), respectively. Thus, cell growth slows down as cells exit mitotic proliferation. Consistent with this, cystocytes become progressively smaller during transit-amplifying divisions (Fig. 2g). *diminutive* (dMyc), the *Drosophila* Myc homologue, an important regulator of cell growth, is highly expressed in stem cells and cystoblasts but downregulated in 16-cell cysts (Fig. 2e). Nucleoli (stained by anti-Fibrillarin)—the sites of ribosomal RNA transcription—are large in stem cells but much smaller in 16-cell cysts (Fig. 2c and Supplementary Fig. 2c). Because ribosome number is thought to control cellular growth in *Drosophila*^{17,18}, a reduction in ribosome biogenesis might be responsible for the reduced cell growth at the end of mitotic proliferation.

In *mei-P26* mutants, cellular and nucleolar size are increased and dMyc is highly expressed throughout the tumour (Fig. 2d, f, g and

Supplementary Fig. 2c): the volume of 16-cell cysts is increased to $3,956 \pm 424 \mu\text{m}^3$ ($n = 4$; Fig. 2b) and cell diameters no longer decrease as cells are displaced from the stem cell niche (Fig. 2g). Thus, *mei-P26* is responsible for the differential regulation of cell growth in the *Drosophila* ovarian stem cell lineage. Like *brat*, it might achieve its function by regulating ribosome biogenesis and controlling the expression of dMyc.

To analyse Mei-P26 expression, we generated a specific antibody (Supplementary Fig. 2d–f). Mei-P26 mRNA and protein levels are low in stem cells but are upregulated in cysts as they decrease growth and exit mitotic proliferation (Fig. 2h, j and Supplementary Fig. 2e, g, h). This regulation is functionally important because *mei-P26* overexpression in ovaries results in stem cell loss and complete depletion of the female germ line (Fig. 3a, b). In *bam* mutant ovaries, Mei-P26 levels remain low suggesting that *mei-P26* is regulated in a Bam-dependent manner (Fig. 2i). *bam* overexpression induces premature differentiation of stem cells in a wild-type (Fig. 3d) but not in a *mei-P26* mutant (Fig. 3e) background, demonstrating that Mei-P26 is essential for Bam to induce cystocyte differentiation. Consistently, germline stem cells do not differentiate when *pumilio* is removed in a *mei-P26* mutant background (Supplementary Fig. 2k, l). To test whether Mei-P26 is the only target of Bam, we overexpressed *mei-P26* in a *bam* mutant background

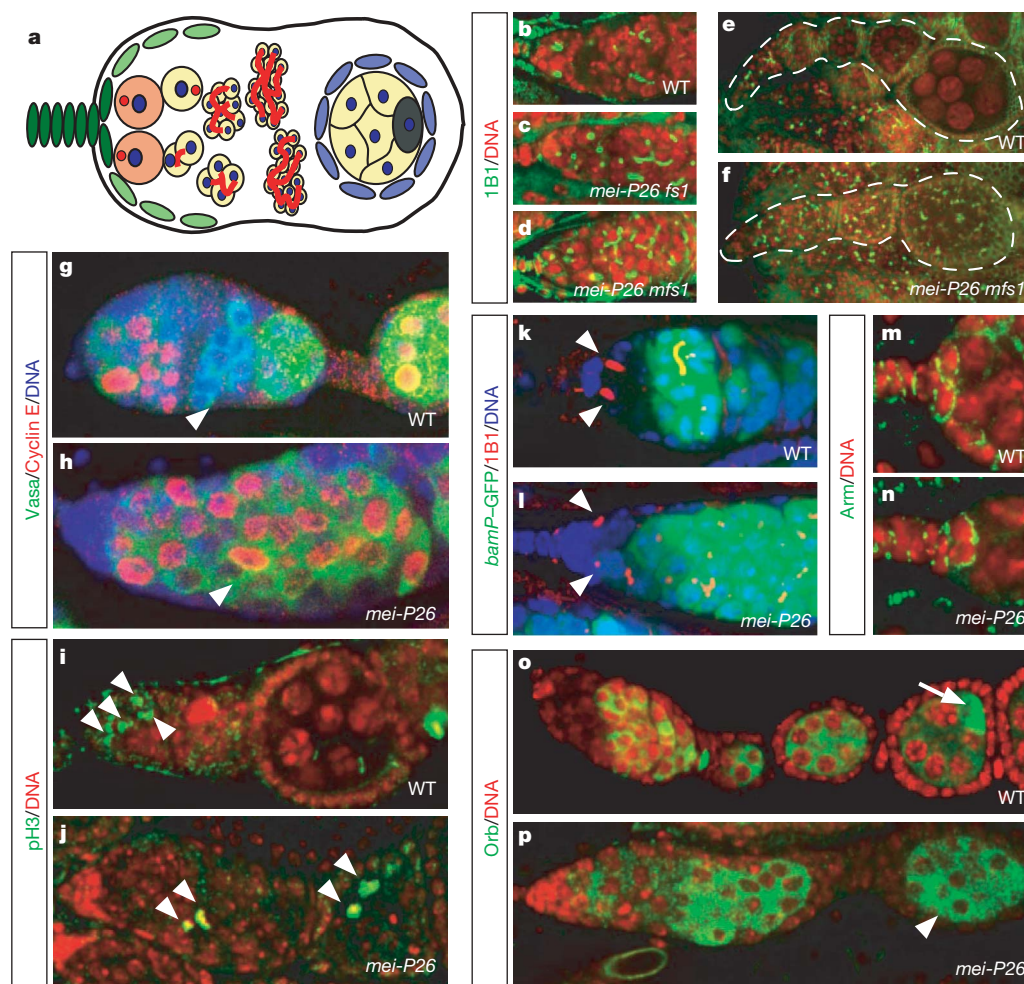


Figure 1 | Differentiation and cell cycle defects in *mei-P26* mutant ovaries. **a**, Overview of *Drosophila* oogenesis showing cap cells (dark green), escort cells (light green), stem cells (orange), cystoblasts and cystocytes (yellow), follicle cells (blue) and the oocyte (dark grey). The spectrosome and fusome are red. **b–f**, Wild type (WT) (**b**, **e**) and *mei-P26* (**c**, **d**, **f**) ovarioles stained with mAb1B1 (1B1, green) and for DNA (red). *fs1* and *mfs1* represent two different *mei-P26* alleles. **g**, **h**, Cyclin E (red; green, Vasa; blue, DNA) is downregulated in WT (**g**, arrowhead) but not in the *mei-P26*^{fs1} mutant (**h**, arrowhead) cystocytes. **i**, **j**, Phospho-H3 positive (green, pH3; red, DNA)

mitotic cells (arrowheads) are restricted to the tip of the ovarioles in WT (**i**) but are found at all stages in *mei-P26*^{fs1} (**j**) ovarioles. **k**, **l**, GFP under the control of the *bam* promoter (*bamP-GFP*) (green; red, mAb1B1; blue, DNA) is not expressed in WT (**k**) and *mei-P26*^{fs1} mutant (**l**) niche contacting germline cells (arrowheads). **m**, **n**, Anti-Armadillo (green; red, DNA) shows integrity of WT (**m**) and *mei-P26*^{fs1} mutant (**n**) cap cells. **o**, **p**, Orb expression is initiated in WT (**o**) and *mei-P26*^{fs1} mutant (**p**) cystocytes but restricted to the oocyte only in WT (**o**, arrow) and not in *mei-P26*^{fs1/mfs1} mutant (**p**, arrowhead) egg chambers.

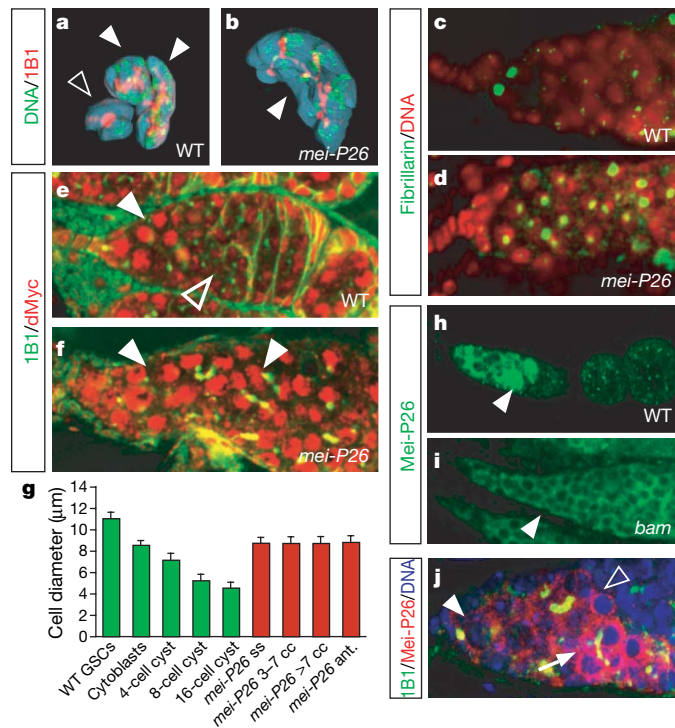


Figure 2 | Mei-P26 regulates cell and nucleolar size. **a, b**, Three-dimensional reconstruction of WT 16-cell cysts (filled arrowheads; **a**), a WT 8-cell cyst (open arrowhead; **a**) and a *mei-P26^{ss1/mfs1}* mutant cyst containing 14 cells (**b**). **c, d**, Nucleoli (green, anti-Fibrillarin) in *mei-P26^{ss1}* mutant ovaries (**d**) are larger than those in the WT ovaries (**c**). **e, f**, High levels of dMyc (red; green, mAb1B1) are detected in germline stem cells and early cysts (filled arrowhead). In postmitotic cysts, dMyc levels are decreased (open arrowhead) and levels increase again as nurse cells undergo endoreplication (**e**). In *mei-P26^{ss1/mfs1}* mutants, high levels of dMyc are detected throughout the tumour (arrowheads in **f**). **g**, Diameter of the indicated cell types in WT ($n > 24$ cells) and *mei-P26^{ss1/mfs1}* mutant ($n = 11$ cells for Mei-P26 ant., $n > 43$ for all others) ovaries. Abbreviations: Mei-P26 ss, single-spectrosome-containing cells; Mei-P26 3–7 cc (or Mei-P26 >7 cc), cystocytes in *mei-P26^{ss1}* mutant cysts containing either 3–7 (or >7) cells; Mei-P26 ant., anterior niche contacting cells. Error bars, s.e.m. (green bars, WT; red bars, *mei-P26* mutant). **h, i**, Mei-P26 expression peaks in early 16-cell cysts in WT and is not detected at later stages of oogenesis (**h**). In *bam^{Δ86}* mutant ovaries, expression is not upregulated (**i**, compare to WT in **h**). Arrowheads point at equivalent stages. Note that Mei-P26 staining appears more intense in later stages owing to sample thickness and out-of-focus fluorescence (**i**). **j**, Germarium close up: Mei-P26 (red; green, mAb1b1; blue, DNA) expression is low in stem cells (arrowhead), weakly upregulated in 8-cell cysts (open arrowhead) and peaks in 16-cell cystocytes (arrow) (see Supplementary Fig. 2e, f, h).

(Fig. 3f, g and Supplementary Fig. 2i, j). *mei-P26* overexpression reduces the size of *bam* mutant cells (from $10.1 \pm 0.17 \mu\text{m}$ (*bam*, $n = 27$) to $7.0 \pm 0.19 \mu\text{m}$ (*bam*, *nanos gal4* >> *UASP-meP26*,

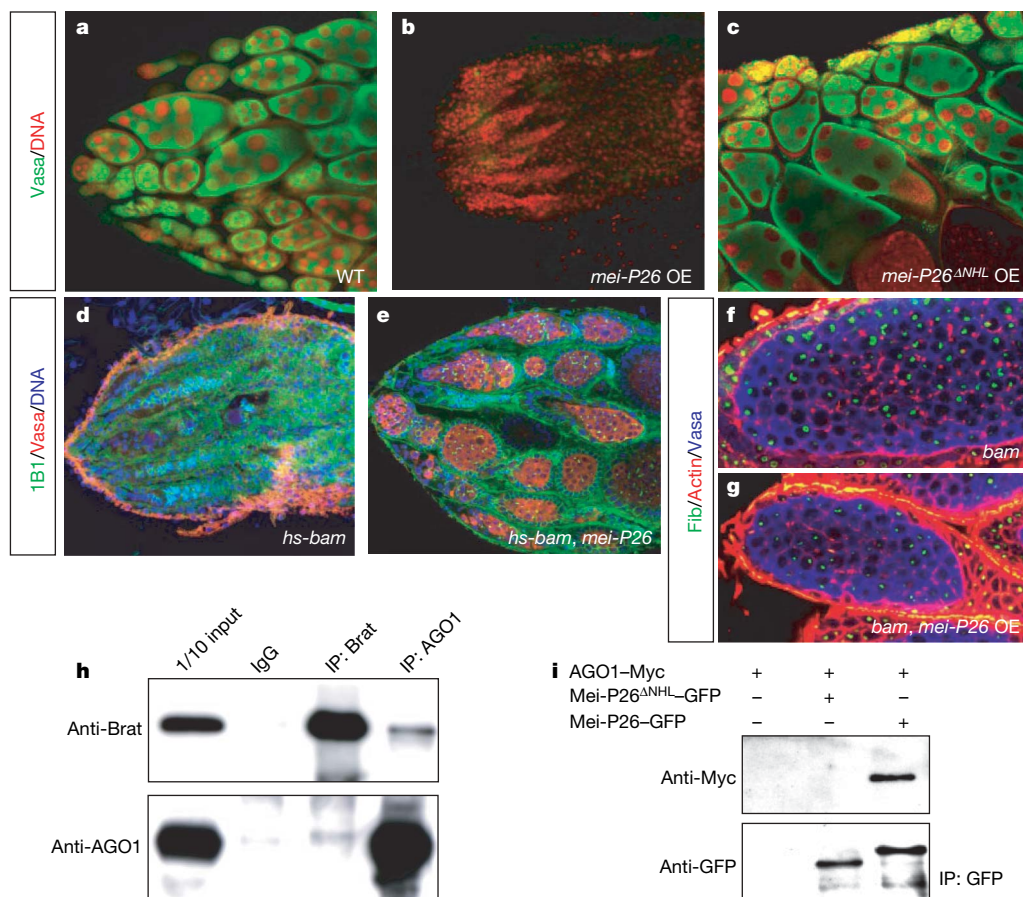


Figure 3 | Bam requires the AGO1-binding protein Mei-P26 to induce proper cystocyte differentiation. **a, b, c**, *mei-P26* (**b**) but not *mei-P26^{ANHL}* (**c**) overexpression (OE) (from *nanos-Gal4::VP16*) depletes the germ line (green, Vasa; red, DNA). **d, e**, Transient (heat-shock-induced) *bam* overexpression induces stem cell differentiation and depletes the germ line (red, Vasa; green, mAb1B1; blue, DNA) in a WT (**d**) but not *mei-P26^{ss1/mfs1}* mutant (**e**) background. **f, g**, *mei-P26* overexpression (from *nanos-Gal4::VP16*) in *bam^{Δ86}* mutants (**g**) reduces cell and nucleolar size (see

statistics in Supplementary Fig. 2c) but does not induce stem cell differentiation (Supplementary Fig. 2i, j). **h, i**, The NHL domain proteins Brat, Mei-P26 and Dappled interact with AGO1. Reciprocal immunoprecipitations (IP) of Brat and AGO1 from *Drosophila* embryo extracts (**h**). GFP-tagged Mei-P26 but not GFP-tagged Mei-P26^{ANHL} (**i**) coimmunoprecipitates Myc-tagged AGO1 in S2 cells (see Supplementary Fig. 4b).

$n = 27$)) and the enlarged nucleolus in *bam* mutants (Supplementary Fig. 2c) but does not rescue the differentiation defects (Supplementary Fig. 2i, j) observed in these mutants. Thus, *mei-P26* is upregulated by Bam activity in cystocytes and inhibits cell growth and mitotic proliferation.

In the brain, *mei-P26* is weakly expressed in neuroblasts, is absent from ganglion mother cells but is highly expressed in neurons (Supplementary Fig. 3a–d). Although *mei-P26* mutants have no obvious defects in neurogenesis, *mei-P26* overexpression can induce premature neuronal differentiation (Supplementary Fig. 3e, f), suggesting that *mei-P26* and *brat* might have some common targets. Using mass spectrometry, we identified the RNase Argonaute-1 (AGO1) in anti-Brat immunoprecipitates (Supplementary Fig. 4a). Brat and AGO1 can be co-immunoprecipitated from *Drosophila* embryos (Fig. 3h), and co-transfection experiments in S2 cells show that AGO1 also binds Mei-P26 and the Trim–NHL protein Dappled (Supplementary Fig. 4b). This interaction is functionally significant because Mei-P26 lacking the NHL domain no longer binds AGO1 (Fig. 3i) and fails to block self renewal when overexpressed in ovarian stem cells (Fig. 3c).

AGO1 is a core component of the RISC complex¹⁹ and is important for microRNA-mediated translational repression and RNA degradation. MicroRNAs are essential for self renewal in *Drosophila* ovarian stem cells because mutations in *AGO1* (Supplementary Fig. 5), *dicer-1* (ref. 20) or its binding partner *loquacious*²¹ result in premature stem cell differentiation. To test whether Mei-P26 regulates microRNAs, we measured microRNA levels by quantitative reverse transcription polymerase chain reaction (RT–PCR). In *mei-P26* mutants, most microRNAs are significantly upregulated (Supplementary Fig. 7a, b), whereas overexpression of *mei-P26* in *bam* mutants broadly reduces microRNA levels (Fig. 4a and Supplementary Fig. 7a). Importantly, the *mei-P26* mutant phenotype can be partially rescued by removing one copy of *Loquacious* and thereby reducing microRNA levels (Fig. 4b, c and Supplementary Fig. 6a–f), indicating that *mei-P26* acts by inhibiting the microRNA pathway in cystocytes.

One of the best characterized *Drosophila* microRNAs is *bantam*, a regulator of proliferation and apoptosis²². We used a sensor that expresses GFP from the *tubulin* promoter and carries two *bantam* binding sites in the 3' untranslated region²². In wild-type ovaries, this sensor is repressed by *bantam* activity in stem cells but is derepressed in 16-cell cystocytes in which high levels of Mei-P26 are present (Fig. 4e). In *mei-P26* mutants, the sensor is off in all germline cells (Fig. 4g). In contrast, a control sensor lacking microRNA-binding sites shows high expression in all wild-type and mutant cells (Fig. 4d, f). Although *mei-P26* regulates many microRNAs, *bantam* seems to be an important target: even animals heterozygous for a *bantam* null mutation have a reduced number of stem cells (1.15 ± 0.12 per gerarium ($n = 32$ germaria) compared to 2.09 ± 0.05 ($n = 32$ germaria) in 14-day-old *bantam* heterozygous and wild-type flies, respectively), suggesting a defect in self renewal (Supplementary Fig. 7e, f). To test whether *mei-P26* regulates *bantam* directly, we expressed a luciferase construct carrying a *bantam*-binding site in its 3' untranslated region in S2 cells. In control cells, the *bantam* sensor is repressed, but a construct lacking the binding site is not affected (Supplementary Fig. 7d). On cotransfection of *mei-P26* (Supplementary Fig. 7c), luciferase expression is significantly derepressed, but no derepression is seen with cotransfection of *mei-P26* lacking the NHL domain; this indicates that Mei-P26 can repress *bantam* activity even in S2 cells.

Our data suggest that *brat* and *mei-P26* might act in a similar manner to control proliferation in stem cell lineages. In both mutants, cells that normally stop self renewal increase ribosome biogenesis, grow abnormally large and fail to exit the cell cycle leading to the formation of a tumour. The general upregulation of microRNAs in *mei-P26* mutants leaves several possibilities for how these proteins might regulate the microRNA pathway. The presence of a RING finger in Mei-P26 suggests a role in protein degradation. The high amounts of AGO1 detected in Mei-P26 immunoprecipitates make it unlikely that AGO1 itself is degraded by Mei-P26. However, another member of the RISC complex might be a degradation target

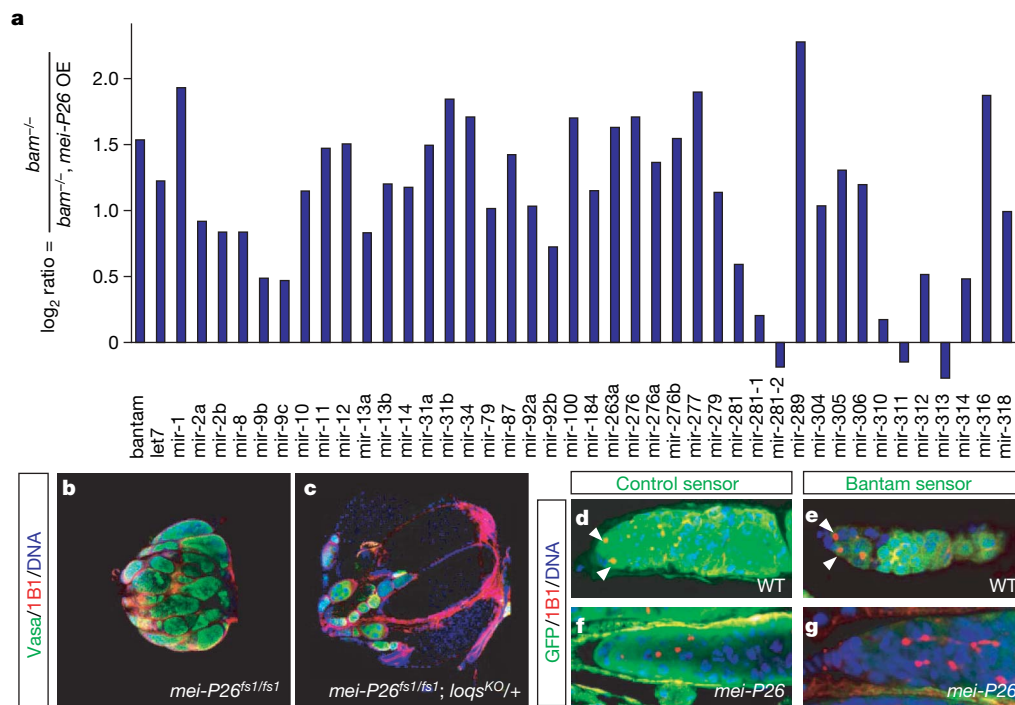


Figure 4 | Mei-P26 regulates miRNAs. **a**, Quantitative PCR experiment comparing the level of mature microRNAs in *bam*^{Δ86} and *bam*^{Δ86}; *nanos-Gal4::VP16* >> *pUASp-meï-P26*. The y axis shows log₂ of the expression ratio between the two genotypes. **b, c**, Loss of one copy of *loquacious* partially rescues the *mei-P26*^{fs1} mutant phenotype (Supplementary Fig. 6).

d–g, Downregulation of the *bantam* sensor reveals *bantam* microRNA expression in stem cells but not cystocytes (e), whereas the control sensor is uniformly expressed throughout the germ line (d). The *bantam* sensor (g) but not the control sensor (f) is silenced in the germ line of *mei-P26*^{fs1/mfs1} mutants (arrowheads indicate stem cells).

of Mei-P26. Equally likely, Mei-P26 could prevent the incorporation or increase the turnover of microRNAs in the RISC complex.

Many human tumours contain cancer stem cells that drive tumour growth and metastasis^{23,24}. Although the similarities between *Drosophila* tumours and human cancer are limited, *brat* mutant brains and *mei-P26* mutant ovaries (as well as the other mutant conditions causing stem cell tumours^{25–27}) provide an invertebrate model for stem-cell-derived tumour formation. In *mei-P26* mutants, tumours originate from cystocytes, the transit-amplifying pool of the ovarian stem cell lineage. In *mei-P26* mutants, these cells re-gain the ability to self-renew: after *bam* overexpression—which leads to premature differentiation of stem cells—the germ line is depleted in a wild-type but not in a *mei-P26* mutant background. Thus, *mei-P26* tumours arise from growth defects in the transit-amplifying compartment of the ovarian stem cell lineage—a mechanism that could occur in human tumours as well.

Our data establish Trim-NHL proteins as regulators of stem cell proliferation. Vertebrate members of this family exist and are down-regulated in human cancer cell lines²⁸ suggesting that their tumour-suppressor function might be conserved in vertebrates as well.

METHODS SUMMARY

For immunofluorescence, ovaries were dissected in PBS, fixed for 20 min in *n*-heptane/PBS (1:1) with 5% PFA, washed three times in PBS containing 0.1% Triton X-100, and stained using standard procedures. Cell volumes were determined after three-dimensional reconstruction from stacks of 0.4- μ m sections using IMARIS software. A Rho1-GFP trap (ZCL1957) was used to outline stem cells. To outline individual cysts, mitotic clones, generated by heat-induced expression of FLP recombinase (*hs-FLP*) and marked by the absence of GFP, were stained with rhodamine-phalloidine. Protein lysates were prepared from 20 ovaries dissected in PBS and extracted with Laemmli buffer, and were run on 10% SDS-polyacrylamide gels. Overexpression constructs were created by amplifying *mei-P26* full length, *mei-P26*^{ANHL}, *AGO1*, *dappled* and *brat* coding sequences by PCR from total complementary DNA or expressed sequence tags (ESTs) with appropriate primers containing bacterial attachment (attB) sites to mediate intermolecular recombination for Gateway cloning (Invitrogen) or appropriate restriction sites. MicroRNAs were quantified using the TaqMan microRNA assay and a quantitative PCR machine (Applied Biosystems) with primer sets obtained from Applied Biosystems on 10 ng total RNA extracted from ovaries with TRIzol Reagent (Invitrogen). To measure *bantam* activity, S2 cells were transfected with a plasmid that expresses the yeast transcription factor Gal4 under the control of the actin promoter (*actin-Gal4*), a control construct expressing *Renilla* luciferase, a sensor carrying two copies of a sequence complementary to the *bantam* miRNA downstream of firefly luciferase (FLuc) and the respective pUASp-Mei-P26 constructs. All transfections were performed in triplicate and dual luciferase assays were performed 24 h after transfection following the manufacturer's instructions (Promega).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 February; accepted 3 April 2008.

Published online 4 June 2008.

- Knoblich, J. A. Mechanisms of asymmetric stem cell division. *Cell* **132**, 583–597 (2008).
- Fuller, M. T. & Spradling, A. C. Male and female *Drosophila* germline stem cells: two versions of immortality. *Science* **316**, 402–404 (2007).
- Gilboa, L. & Lehmann, R. How different is Venus from Mars? The genetics of germ-line stem cells in *Drosophila* females and males. *Development* **131**, 4895–4905 (2004).
- Lee, C. Y., Wilkinson, B. D., Siegrist, S. E., Wharton, R. P. & Doe, C. Q. Brat is a Miranda cargo protein that promotes neuronal differentiation and inhibits neuroblast self-renewal. *Dev. Cell* **10**, 441–449 (2006).
- Bello, B., Reichert, H. & Hirth, F. The brain tumor gene negatively regulates neural progenitor cell proliferation in the larval central brain of *Drosophila*. *Development* **133**, 2639–2648 (2006).
- Betschinger, J., Mechtler, K. & Knoblich, J. A. Asymmetric segregation of the tumor suppressor brat regulates self-renewal in *Drosophila* neural stem cells. *Cell* **124**, 1241–1253 (2006).
- Sekelsky, J. J. *et al.* Identification of novel *Drosophila* meiotic genes recovered in a P-element screen. *Genetics* **152**, 529–542 (1999).
- Page, S. L., McKim, K. S., Deneen, B., Van Hook, T. L. & Hawley, R. S. Genetic studies of mei-P26 reveal a link between the processes that control germ cell proliferation in both sexes and those that control meiotic exchange in *Drosophila*. *Genetics* **155**, 1757–1772 (2000).
- Reymond, A. *et al.* The tripartite motif family identifies cell compartments. *EMBO J.* **20**, 2140–2151 (2001).
- Zaccai, M. & Lipshitz, H. D. Differential distributions of two adducin-like protein isoforms in the *Drosophila* ovary and early embryo. *Zygote* **4**, 159–166 (1996).
- Lin, H., Yue, L. & Spradling, A. C. The *Drosophila* fusome, a germline-specific organelle, contains membrane skeletal proteins and functions in cyst formation. *Development* **120**, 947–956 (1994).
- Ohlmeyer, J. T. & Schubach, T. Encore facilitates SCF-ubiquitin-proteasome-dependent proteolysis during *Drosophila* oogenesis. *Development* **130**, 6339–6349 (2003).
- Chen, D. & McKearin, D. M. A discrete transcriptional silencer in the bam gene determines asymmetric division of the *Drosophila* germline stem cell. *Development* **130**, 1159–1170 (2003).
- Lantz, V., Chang, J. S., Horabin, J. I., Bopp, D. & Schedl, P. The *Drosophila* orb RNA-binding protein is required for the formation of the egg chamber and establishment of polarity. *Genes Dev.* **8**, 598–613 (1994).
- Kai, T. & Spradling, A. Differentiating germ cells can revert into functional stem cells in *Drosophila melanogaster* ovaries. *Nature* **428**, 564–569 (2004).
- Frank, D. J., Edgar, B. A. & Roth, M. B. The *Drosophila melanogaster* gene brain tumor negatively regulates cell growth and ribosomal RNA synthesis. *Development* **129**, 399–407 (2002).
- Grewal, S. S., Li, L., Orian, A., Eisenman, R. N. & Edgar, B. A. Myc-dependent regulation of ribosomal RNA synthesis during *Drosophila* development. *Nature Cell Biol.* **7**, 295–302 (2005).
- Rudra, D. & Warner, J. R. What better measure than ribosome synthesis? *Genes Dev.* **18**, 2431–2436 (2004).
- Tolia, N. H. & Joshua-Tor, L. Slicer and the argonautes. *Nature Chem. Biol.* **3**, 36–43 (2007).
- Jin, Z. & Xie, T. Dcr-1 maintains *Drosophila* ovarian stem cells. *Curr. Biol.* **17**, 539–544 (2007).
- Park, J. K., Liu, X., Strauss, T. J., McKearin, D. M. & Liu, Q. The miRNA pathway intrinsically controls self-renewal of *Drosophila* germline stem cells. *Curr. Biol.* **17**, 533–538 (2007).
- Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B. & Cohen, S. M. *bantam* encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in *Drosophila*. *Cell* **113**, 25–36 (2003).
- Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111 (2001).
- Clarke, M. F. & Fuller, M. Stem cells and cancer: two faces of eve. *Cell* **124**, 1111–1115 (2006).
- Causinus, E. & Gonzalez, C. Induction of tumor growth by altered stem-cell asymmetric division in *Drosophila melanogaster*. *Nature Genet.* **37**, 1125–1129 (2005).
- Lee, C. Y. *et al.* *Drosophila* Aurora-A kinase inhibits neuroblast self-renewal by regulating aPKC/Numb cortical polarity and spindle orientation. *Genes Dev.* **20**, 3464–3474 (2006).
- Wang, H. *et al.* Aurora-A acts as a tumor suppressor and regulates self-renewal of *Drosophila* neuroblasts. *Genes Dev.* **20**, 3453–3463 (2006).
- Ross, D. T. *et al.* Systematic variation in gene expression patterns in human cancer cell lines. *Nature Genet.* **24**, 227–235 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank V. Siegel, K. Mochizuki, G. B. Cebolla and S. Weitzer for comments on the manuscript, J. Stolte for assistance with quantitative PCRs, C. Richter and the other members of the Knoblich laboratory for discussion, B. Dickson, E. Izaurralde, P. Lasko, L. Luo, S. Hawley, D. McKearin, H. Richardson, F. Schnorrer, J. Skeath, D. Stein, L. Wong, the Bloomington *Drosophila* Stock Center, the Developmental Studies Hybridoma Bank (DSHB) and the *Drosophila* Genomics Resource Center (DGRC) for reagents, and M. Insko and M. Fuller for communicating results before publication. Work in the Knoblich laboratory is supported by the Austrian Academy of Sciences, the Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF), the Austrian Science Fund (FWF) and the EU network ONCASYM; K.M. is supported by the Austrian Proteomics Platform (APP) of the Austrian Genome Program (GENAU).

Author Contributions J.A.K. and R.A.N. designed the study. R.A.N. performed the oogenesis experiments. J.B. and K.M. contributed biochemical data (Fig. 3h and Supplementary Fig. 4a, b). A.F. contributed the S2 luciferase assay (Supplementary Fig. 7c, d). I.P. assisted in the experiment in Fig. 3f, g and performed experiments in the larval brain (Supplementary Fig. 3). S.C. and N.B. designed the microRNA quantitative PCR experiment, which was performed by N.B. (Fig. 4a and Supplementary Fig. 7a, b). J.A.K. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.A.K. (juergen.knoblich@imba.oeaw.ac.at).

METHODS

Cytology and immunofluorescence. Immunofluorescence experiments in larval brains were carried out as described previously⁶. Ovaries were dissected in PBS and fixed for 20 min in a 1:1 mixture of *n*-heptane and PBS containing 5% PFA. After three washes in PBS containing 0.1% Triton X-100, they were treated as described for brains⁶. For *in situ* hybridizations, DIG-labelled antisense probes were synthesized from the EST GH01646. For clonal analysis, Flp expression was induced by incubating flies at 37 °C for 1 h. *hs-bam* (Fig. 3d, e) was induced by three consecutive 1 h incubations at 37 °C within one day during development. Overexpression of *mei-P26* in the larval brain was carried out with a temperature-sensitive Gal80 (a yeast protein that suppresses Gal4) to circumvent embryonic lethality, and expression was induced for 60 h at 29 °C. The following antibodies were used: mouse anti-1B1 (7H9, Developmental Studies Hybridoma Bank (DSHB), 1:1), rabbit anti-AGO1 (Abcam, 1:100), mouse anti-Armadillo (N2 7A1, DSHB, 1:100), rabbit anti-Brat (1:100), mouse anti-CycE (from H. Richardson, 1:10), guinea pig anti-Deadpan (gift from J. Skeath, unpublished, 1:1,000), mouse anti-Fibrillarin (Abcam, 1:10), mouse anti-GFP (Roche, 1:100), rabbit anti-GFP (Abcam, 1:100), mouse anti-c-Myc (Santa Cruz Biotechnology, 1:100), rabbit anti-d-Myc (from D. Stein 1:5,000), rabbit anti-phosphorylated histone H3 (Upstate Biotechnology, 1:1,000), mouse anti-Prospero (DSHB, 1:10), mouse anti-Orb (4H8, DSHB, 1:10), mouse anti- α -Tubulin (Sigma), rabbit anti-Vasa (gift from P. Lasko, 1:10,000) and goat anti-Vasa (Santa Cruz, 1:200). Rhodamine-conjugated phalloidine (Alexa) was used 1:1,000 and rabbit anti-Mei-P26 (1:300) was raised against the peptide NLKTVLSDDASNSSVLED corresponding to amino acids 23–40. Samples were mounted in Vectashield containing 4,6-diamidino-2-phenylindole (DAPI; Vector Laboratories) and imaged on a Zeiss LSM 500 confocal microscope equipped with a blue diode laser to visualize DAPI. Images were processed in Adobe Photoshop and Adobe Illustrator.

Constructs. The *mei-P26* full-length, *mei-P26^{ANHL}* (amino acids 1–933), *AGO1*, *dappled* and *brat* coding sequences were PCR-amplified from total complementary DNA or ESTs with appropriate primers containing attB recombination sites (or with primers containing appropriate restriction sites), and were sequenced and recombined into pDONR221 (Invitrogen) and thereafter recombined into either tagged or untagged pUASp or pUAST destination vectors (DGRC). For fly transformation, constructs were co-injected with $\Delta 2$ –3 transposase into *Drosophila w¹¹¹⁸* embryos.

To generate a *bantam* luciferase reporter, two copies of a sequence complementary to the *bantam* microRNA were cloned downstream of the FLuc reporter plasmid (*bantam* sensor). A plasmid expressing *Renilla* luciferase (RLuc) (gift from E. Izaurralde) served as a transfection control.

Biochemistry and S2 cell experiments. For ovary lysates, 20 ovaries from the respective genotypes were dissected in PBS. Proteins were extracted with Laemmli

buffer and run on a 10% SDS-page. S2 cell transfections, immunoprecipitations, silverstaining and mass spectrometry were performed essentially as described⁶.

For microRNA assays, S2 cells were simultaneously transfected with Actin–Gal4, the FLuc microRNA reporter construct, the RLuc transfection control construct and pUASp *mei-P26*, pUASp *mei-P26^{ANHL}* or the empty pUASp vector as a control. Dual luciferase assays were performed 24 h after transfection following the manufacturer's instructions (Promega). For each experiment, transfections were performed in triplicates. Three independent experiments showed comparable results, one of which is shown in Supplementary Fig. 7.

Fly strains. We used the following *Drosophila* strains: *AGO1^{k08121}* (gift from T. Uemura), *bantam^{A1}*, *mei-P26^{fs1}* (ref. 8), *mei-P26^{mfs1}* (ref. 8), *bam^{A86}* (Bloomington), *loqs⁰⁰⁷⁹¹* (Bloomington), *loqs^{KO}* (gift from D. McKearin), *nanos-Gal4::VP16* (gift from F. Schnorrer), *hs-Bam* (gift from D. McKearin), *brat⁹²*, *FRT18A Ubi-GFP.nls*, *hs-Flp/CyO* (Bloomington), *yw hsFLP; FRT G13 2xGFP.nls*, *UASp-me-P26* (gift from S. Hawley, unpublished), *BamP-GFP* (gift from D. McKearin), the *Gal4* lines *OK107* (gift from L. Luo) and *1407*, *tub-Gal80^{ts}* (Bloomington), *Rho1-GFP trap* (ZCL1957, Fly Trap), and *pum⁰¹⁶⁸⁸* (Bloomington), *pum^{ET3}*. All fly strains were raised on standard food without wet or dry yeast.

microRNA quantitative PCRs. Total RNA from ovaries was extracted from the four respective genotypes (wild type, *mei-P26^{mfs1/fs1}*, *bam^{A86}* and *bam^{A86} nanos-Gal4::VP16* \gg *pUASp-me-P26*) using the TRIzol Reagent (Invitrogen). Primer sets designed to amplify mature microRNAs (and sno RNA227 as a control reaction) were obtained from Applied Biosystems. Products were amplified from 10 ng total RNA samples from the respective genotypes with the TaqMan microRNA assay using a quantitative PCR machine (Applied Biosystems). Wild-type microRNA levels are set to zero in Supplementary Fig. 7a.

Quantification. Wild-type and *mei-P26^{mfs1}* flies carrying *BamP-GFP* were stained with mAb1B1 to identify cell types and rhodamine-phalloidine to outline cells. Cell diameters were determined from image stacks of 0.4- μ m sections using the measuring tool of LSM software (Zeiss). Cell types were defined as follows: Stem cells: GFP-negative, niche-contacting single-spectrosome-containing cells. Cystoblasts: GFP-positive single-spectrosome-containing cells. Cystocytes: 4, 8 or 16 GFP-positive cells connected by fusomes. To determine nucleolar:cell size ratio, anti-Fibrillarin and rhodamine-phalloidine were used in stainings. For ellipsoid cells, the mean cell diameter was determined (from the longest and shortest cell diameter) and used in the statistics.

For volume reconstruction of germline stem cells, *Rho1-GFP trap* (ZCL1957) flies were stained with mouse anti-1B1, GFP and DAPI, and volumes were determined from image stacks of 0.4- μ m sections using the IMARIS software. For reconstruction of cyst volume, GFP-negative clones, stained with mouse anti-1B1, rhodamine-conjugated phalloidine (which perfectly colocalizes with the *Rho1-GFP trap* localization pattern within the resolution of our microscope), GFP and DAPI and were analysed in a similar manner.

CORRIGENDUM

doi:10.1038/nature07133

Neurophysiology: Sensing temperature without ion channels

Brandon R. Brown

Nature 421, 495 (2003)

My Brief Communication about thermoelectricity in shark gels neglected a systematic effect of surface electrochemistry: electrode potentials vary with temperature in electrolyte solutions. However, silver leads in sea water¹ and accepted values for likely electrode reactions² show a sign opposing the gel signals, making it unlikely that an artefactual signal is the origin. Our subsequent work³ discussed artefacts and repeated the signal with platinum electrodes. Although another report⁴ finds a zero signal using salt bridges, it ignores thermopower in gel-filled leads, which risks building a 'null thermocouple' from two similar materials (see ref. 5, for example). A temperature function of the electrosensors is not known, but the thermoelectric transduction hypothesis still stands.

1. Sanford, T. B. *Measurements and Interpretations of Motional Electric Fields in the Sea*. PhD thesis, Massachusetts Institute of Technology (1967).
2. Milazzo, G. & Caroli, S. *Tables of Standard Electrochemical Potentials* (John Wiley and Sons, New York, 1978).
3. Brown, B. R., Hughes, M. E. & Russo, C. Thermoelectricity in natural and synthetic hydrogels. *Phys. Rev. E* 70, 031917 (2004).
4. Fields, R. D., Fields, K. D. & Fields, M. C. Semiconductor gel in shark sense organs? *Neurosci. Lett.* 426, 166–170 (2007).
5. Kasap, S. O. *Principles of Electronic Materials and Devices* 278–284 (McGraw Hill, San Francisco, 2000).

CORRIGENDUM

doi:10.1038/nature07134

An extremely luminous X-ray outburst at the birth of a supernova

A. M. Soderberg, E. Berger, K. L. Page, P. Schady, J. Parrent, D. Pooley, X.-Y. Wang, E. O. Ofek, A. Cucchiara, A. Rau, E. Waxman, J. D. Simon, D. C.-J. Bock, P. A. Milne, M. J. Page, J. C. Barentine, S. D. Barthelmy, A. P. Beardmore, M. F. Bietenholz, P. Brown, A. Burrows, D. N. Burrows, G. Bryngelson, S. B. Cenko, P. Chandra, J. R. Cummings, D. B. Fox, A. Gal-Yam, N. Gehrels, S. Immler, M. Kasliwal, A. K. H. Kong, H. A. Krimm, S. R. Kulkarni, T. J. Maccarone, P. Mészáros, E. Nakar, P. T. O'Brien, R. A. Overzier, M. de Pasquale, J. Racusin, N. Rea & D. G. York

Nature 453, 469–474 (2008)

In this Article, the surname of co-author G. Bryngelson was misspelled as G. Byrngelson.

naturejobs

**THE CAREERS
MAGAZINE FOR
SCIENTISTS**

Music emanated from the aircraft carrier, as thousands waited in line, eager to get on deck. Already the party was in full swing. But this wasn't a political gathering or a military celebration. This was a gala for the Biotechnology Industry Organization, part of the group's annual international conference. Held in mid-June, the gathering took place on the *USS Midway*, a retired aircraft carrier that has been converted into a museum and is docked near downtown San Diego. After a day of sessions, thousands of people from biotech workers and venture capitalists to technology-transfer officers and journalists gathered on a warm, breezy California evening to indulge in free food, drink and entertainment.

Conferences serve many purposes — to vet ideas among peers, to seek new collaborators and employers, to learn and, yes, to socialize. Indeed, sometimes the networking and the socializing are quite intertwined. Attendees listened to the two live bands, ate Asian-inspired cuisine and took free rides on the flight simulators. Fireworks from across the marina lit up the night sky. I marvelled at the large Polynesian statues on deck, apparently hoisted up by cargo lift. Clearly the meeting was about displaying and attracting money and investment, not just discussing science.

Naturejobs has made its own slightly less lavish foray into the conference business. Last year we held our inaugural Source Event, a careers forum to advise young scientists on the paths to industry, academia and non-traditional careers. We are building on that success with a larger conference on 26 September in London. Already we have reached half our capacity, with more than 400 registrants. Anyone interested in joining us can get the benefit of some of our panelists' expertise online right now through our Source Event Nature Network group at <http://network.nature.com/group/thesourceevent>. Both online and in person, we can promise sage careers advice, from public- and private-sector researchers, that's tailored to your interests and curiosities. But we can't promise flight-simulator rides — at least not this year.

Gene Russo is editor of *Naturejobs*.

CONTACTS

Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
Business Development Manager:
Amelie Pequignot (4974)
e-mail: a.pequignot@nature.com
Natureevents:

Claudia Paulsen Young (+44 (0) 20 7014 4015)
e-mail: c.paulsenyoung@nature.com
France/Switzerland/Belgium:
Muriel Lestringuez (4994)
Southwest UK/RoW: Nils Moeller (4953)

Scandinavia/Spain/Portugal/Italy:
Evelina Rubio-Hakansson (4973)
Northeast UK/Ireland:
Matthew Ward (+44 (0) 20 7014 4059)
North Germany/The Netherlands:
Reya Silao (4970)
South Germany/Austria:
Hildi Rowland (+44 (0) 20 7014 4084)

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com
Naturejobs web development: Tom Hancock
Naturejobs online production: Dennis Chu

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718

Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

India
Vikas Chawla (+91 1242881057)
e-mail: v.chawla@nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe (+81 3 3267 8765)
e-mail: a.watanabe@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To (+852 2811 7191)
e-mail: g.to@natureasia.com



CATALONIAN POWERHOUSE

Approaching Barcelona by plane over the azure Mediterranean Sea offers breathtaking views. The city's beachfront and landmark skyscrapers reveal why the proud capital of Catalonia has become a major tourist destination. And in the midst of that vista, the omega-shaped complex of the Barcelona Biomedical Research Park (BBRP) not only complements the city's architectural style, but signifies the city's growing ambitions in science.

Spanish science stagnated under a 40-year right-wing dictatorship followed by 30 years of, at best, timid reforms. But from 2004, the newly elected Socialist government increased science funding by 60% and started tackling the bureaucracies that obstructed change and froze out new young talent (see *Nature* 451, 1029; 2008). The government emphasized its commitment after re-election earlier this year by creating a science and innovation ministry, led by biologist Cristina Garmendia.

Barcelona in the lead

The Autonomous Community of Catalonia, always one of the most industrialized regions of Spain, was way ahead. The national funding windfall expanded the regional government's science initiative, launched in 2003, which set up 30 new institutes. These institutes are all public-private foundations, mostly covering aspects of the life sciences, and nearly all are located in and around Barcelona. Catalonia was also ahead of the game when it tackled the laws that promoted 'inbreeding' — which used to allow universities to favour local and internal candidates, leaving little room for newcomers or young people. Catalonia's financial contribution to science has grown dramatically since 2003, spending €2 billion (US\$4.7 billion) during 2005–08.

The BBRP, for example, is just one of Barcelona's three thriving science parks. A joint initiative of the

Spain is revitalizing its science base, with Barcelona surging ahead as a Mediterranean science hub, reports **Quirin Schiermeier**.



The Barcelona Biomedical Research Park attests to the city's ambitions in science.

Catalan government, the city council of Barcelona and Pompeu Fabra University, it opened two years ago. It now houses the Centre for Genomic Regulation (CRG) and five other public research institutes. Some 80 BBRP groups study a broad variety of topics, from stem cells to computational biology. And in a way that could hardly have been imagined 10 years ago, all these groups and institutes are flourishing with international personnel and collaborations, aided by increased salaries and a loosening of traditionally restrictive Spanish recruitment policies.

The CRG now has 28 groups with 10 more teams yet to be recruited, eventually rising to some 400 scientists. It is modelled on the lines of the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, with which it now has a systems-biology collaboration. Like EMBL, the CRG encourages movement by giving group leaders five-year contracts with a maximum four-year extension. It always has an eye open for talented PhD students and postdocs from Europe and elsewhere.

"In the past five years, Barcelona has turned into a premier location for biomedical research," says Luis Serrano, the former head of EMBL's structural-biology programme, who is now in charge of the CRG's systems-biology unit. He says the city has a "California spirit" — an attractive location and a strong international science base — that has helped attract scientists and technicians to the centre's core genomics and proteomics facilities.

The city's other science parks are thriving too. The Barcelona Science Park, next to the University of Barcelona, opened in 2003 and has about 1,000 scientists, mostly in the life sciences. Its Institute for Research in Biomedicine has just set up an oncology programme, for which it is now recruiting group leaders and researchers. The Autonomous University of Barcelona's Science and Technology Park, which

officially opened in 2007, houses research institutes and companies in fields ranging from agro-technology, health and life sciences to physics. At full capacity, it will house more than 1,000 scientists and staff.

Computational power is also increasing. The Barcelona Supercomputing Center houses one of the world's most powerful computers — the MareNostrum, which can do 94.21 trillion operations per second — in what used to be the chapel of the University of Barcelona (pictured opposite). The computer is used for a range of research including protein dynamics, drug testing and modelling in Earth sciences. Meanwhile, the €120-million ALBA Particle Accelerator, funded jointly by national and Catalan governments, is being built near Bellaterra, just outside Barcelona. When it comes online in 2010 it will be deployed in areas such as structural biology and materials research and will have a staff of 100, including technicians, researchers and graduate students.

Wanderers welcome to return

One major advantage Catalonia has over other Spanish regions is the ease with which institutions such as the CRG can hire talent from other countries. This results from initiatives put into place by Andreu Mas-Colell, an economist who served as Catalan minister for universities and research from 2000 until 2003. Traditionally in Spain, as in several other European countries, professors have been appointed by the education ministry in charge, not by the university department. (Universities were recently placed within the country's new Ministry of Science and Innovation; see page 147.) Salaries are kept at modest levels, making it hard to compete for talent. Often Spanish universities do not even recognize foreign degrees; they may require a candidate — even one with a degree from a prestigious institution — to take additional tests. This often means Spanish universities hire natives first; it also means that Spanish scientists who go abroad for postdocs can have a hard time returning. Mas-Colell's framework for the 30 new institutes enabled them to hire internationally, without having to ask the government for permission.

"Things have indeed developed very well," says Mas-Colell, who will next year take over as president of the European Research Council (ERC) in Brussels. He notes that some initiatives first tried in Catalonia are now being copied in the Basque country and elsewhere in Spain.

Like its predecessors, the current regional government coalition, led by the Catalan Socialist Party, has put money and thought into supporting science. But the past few years' growth rates of Spanish science budgets — on both national and state level — cannot be maintained forever, says Mas-Colell. He is confident, however, that the new centres are already maturing: "I hope that the system is already strong enough for maintaining in the future what we have achieved in the recent past."

Public perception and support are important, given the extra money being spent. "Spanish society has clearly changed its mind about science," says Lluís Torner, director of the Institute for Photonic Sciences (ICFO) in Barcelona, suggesting that the country has

shaken off its previous apathy. Like the CRG and other new centres, the ICFO has rapidly expanded since its creation in 2003. Its 15 groups — including around 60 postdocs, 80 graduate students and senior scientists from Germany, the Netherlands, France, Britain and Spain — research photonics for applications in biomedical and environmental sciences. Optical devices developed at the ICFO include tools for medical imaging, environmental sensors for pollution, and research into quantum computing and solar energy. A spin-off company, Radiantis, set up in 2005, markets advanced tuneable and wavelength-conversion laser systems for a variety of applications including spectroscopy, photochemistry and quantum optics. By 2011, the ICFO will have some 300 researchers, according to Torner. "We have

several openings for group leaders, and we can offer excellent conditions to young scientists who wish to become independent and run their own group," he says. The attractiveness of Barcelona as a city of science, culture and sandy beaches does help, he adds. "But we do know, of course, that our competitors in Munich, Amsterdam or Paris can offer good conditions as well."

Gathering grant-winners

Another sign of Catalonia's success is the number of ERC starting grants awarded to young investigators based in and around Barcelona. Of the 24 Spain-based scientists who have won one of the 300 or so grants awarded in the first round earlier this year, 15 are working in Catalonia. In fact, Catalonia has been one of the most successful regions in Europe for ERC awards. "You need someone to bet on you," says Joan Seoane, a cancer researcher at Barcelona's Vall d'Hebron Institute of Oncology, who was awarded a five-year, €1.5-million ERC starting grant for his research on molecular mechanisms involved in brain-tumour formation. "This grant gives me enough security to try something a bit more ambitious"

Seoane's return to his native Barcelona in 2004, after five years at the Memorial Sloan-Kettering Cancer Center in New York, was supported by the Catalan Institution for Research and Advanced Studies, a foundation set up in 2001 to consolidate a group of high-level scientists of all origins. The institution has so far hired some 200 high-profile researchers on a permanent basis.

"There's just a lot of effort in Catalonia to attract people from abroad," says Seoane's wife Luisa Maria Lois, a plant molecular biologist at Barcelona's Centre for Research in Agricultural Genomics. "That's very good for us because it allowed us to come back."

In a welcome coincidence, Barcelona-born Lois, who had been at New York's Rockefeller University, learned just days after her husband that she too had won an ERC grant: a five-year, €1.1-million grant for her research on protein regulation of plant responses to environmental stress. "Returning to Barcelona was something we thought would be difficult," she says. "Now this grant has totally changed my perspective, allowing me to build up a competitive research team. That's more than national agencies can offer."

Quirin Schiermeier is Nature's Germany correspondent.

"In the past five years, Barcelona has turned into a premier location for biomedical research."

— Luis Serrano



Joan Seoane (top), Luisa Maria Lois: returning to Barcelona on ERC awards.

MOVERS

James Ryan, founding dean, Joint School of Nanoscience and Nanoengineering, North Carolina A&T State University and the University of North Carolina at Greensboro



2005-08: Associate vice-president of technology and professor of nanoscience, University at Albany, Albany, New York

2002-05: Distinguished engineer and director of advanced materials and process technology, IBM-Albany NanoTech Center, Albany, New York

When James Ryan received his PhD in chemistry in 1988, nanotechnology wasn't yet a fashionable field nor, for that matter, a pervasive buzzword. Now, building a nanoscience school is his primary career mission. As founding dean of the new Joint School of Nanoscience and Nanoengineering of the University of North Carolina at Greensboro and North Carolina A&T State University, Ryan hopes to use his industry background to make it a world-class institution.

Although Ryan spent much of his career in industry, he knew he would return to academia. He earned his PhD in chemistry at Rensselaer Polytechnic Institute in Troy, New York, as well as a master's in biomedical engineering. Ryan soon accepted a position at IBM, and decided to find a dissertation topic while working at the world's largest computer company. His focus would become the chemistry and materials science of thin films used to make integrated circuits.

When Ryan began his career, he etched features in silicon nitride films that were 6 micrometres (or 6,000 nanometres) wide. The last programme he worked on etched features 90 nanometres wide. "I got a chance to contribute to improvements in semiconductor-chip cost and performance that have enabled the technology advances we have today in computers, communications, defence, health care and entertainment," he says. Ryan contributed to 47 patents in materials science and devices, and received the coveted title of distinguished engineer at IBM, the pinnacle of engineering recognition there.

Colleague Ronald Goldblatt says Ryan assembled the political will to develop IBM's partnership with academia — the IBM-Albany NanoTech Center for Semiconductor Research. "The Albany Center would not exist without Jim Ryan," says Goldblatt. Ryan then moved to academia, as associate vice-president for technology and nanoscience professor at the University at Albany in New York.

After three years, he jumped at the opportunity to build the Joint School of Nanoscience and Nanoengineering. He says corporate partnerships will not only encourage regional economic growth, but also help the school quickly establish novel research and educational opportunities.

In that regard, Ryan sees his industry background as a bonus for the nascent field. "My résumé looks different from my academic peers'," he says. "I have more patents than I do publications. But I know how to effectively combine strengths and needs between organizations." ■
Virginia Gewin

BRICKS & MORTAR

Top states vie to fund more science

Two US states are earmarking substantial funds in a bid to cement their position as leaders in research, infrastructure and talent. On 16 June, Massachusetts signed a \$1-billion science package into law. The same day, Maryland governor Martin O'Malley put forward a similar proposal worth \$1.1 billion.

The money will mean new research institutes, biotech companies and jobs for Massachusetts — and for Maryland, if the legislature approves O'Malley's proposal. Massachusetts' programme includes \$500 million for infrastructure projects; \$250 million for loans, grants and biotech investment; and \$250 million for life-science project funding. Dan O'Connell, Massachusetts secretary of housing and economic affairs, says the initiative will generate about 250,000 jobs over 10 years.

In Maryland, O'Malley's proposed investment would potentially double the number of jobs over the next 10 to 20 years, says David Edgerley, the state's secretary of business and economic development. The plan includes \$60 million to build biotech incubators, \$100 million for other life-sciences facilities and \$200 million for stem-cell research.

The two states already rank top in the United States for overall science

and technology funding, according to the Milken Institute's State Technology and Science Index, released last month. Massachusetts is first, and Maryland jumped from the number four spot in 2004 to the runner-up position in 2007.

US states are increasingly tapping into their own tax bases, as federal funding for R&D has been flat for several years. But legislation for long-term state funding has not always ensured a stream of research dollars; other initiatives, such as Michigan's use of a tobacco settlement (see *Nature* **423**, 203; 2003) have been sidetracked as states work to balance their budgets.

Maryland's investments would drive job creation and contribute to the state's tax base, Edgerley says. He points to Maryland's 4.2 hectares of incubators, which already contribute more than 14,000 jobs. Increasing incubator space by 50% should boost that number considerably, he says. The rankings and proposals have kindled a friendly rivalry between the two states. "They're proposing it," says O'Connell of Maryland's billion-dollar plan. "We've done it." O'Connell notes that it took Massachusetts a year to go from proposal to law. Maryland's legislature will vote on the bill this autumn. ■
Paul Smaglik

POSTDOC JOURNAL

I'll take one 'eureka' please

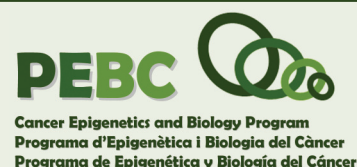
I haven't exclaimed "eureka!" in a long time. In fact, now that I think about it, I never have. Yet I know those moments exist, because I've read about them in books such as James Watson's *The Double Helix*, the story of how the structure of DNA was discovered. And I've been privy to conversations with a few seasoned colleagues in which they recounted their biggest discoveries.

So why haven't I had a eureka moment, I wonder? Is it possible that my projects hold little promise for major discoveries? Or maybe it's because I'm too young, and eureka moments only arrive when researchers are more scientifically mature. I know I have a eureka moment in me. I know it because I'm capable of equally emotional revelations — like recently, when I used a few preferred expletives on hearing that my manuscript had been rejected in less than four hours.

I could use a eureka moment, but I know that science isn't necessarily about expecting or striving for one. I look at my research as a marathon, with experimental ups and downs, unexpected right and left turns, and a continuous process of learning. That may sound trite, but it's a mindset that helped me survive the rigours of being a graduate student, a postdoc and, I hope soon, a professor. Still, some day it would be nice to say eureka, even in a low whisper. ■
Zachary Lippman is a postdoctoral fellow at the Hebrew University of Jerusalem's faculty of agriculture.



Av. Gran Via, s/n Km 2,7
08907 L'Hospitalet - Barcelona
Tels. 93 260 77 33 / 93 335 70 11
Fax 93 260 77 83
www.iconcologia.net



Positions at the Cancer Epigenetics and Biology Program (PEBC), Barcelona

The Catalan Institute of Oncology (Institut Català d'Oncologia, ICO) is a nonprofitable public company, created in March 1995 and member of the Servei Català de la Salut (Catalan Health Service). The Institute is a centre of integral oncology assistance, which, according to international tendency, integrates in only one organization all the elements and the efforts to fight effectively and efficiently against cancer. The Institute is a centre highly specialized and advanced in oncology which offers in a coordinated way: diagnostic, treatment, prevention, training and research. In accordance with the comprehensive cancer centre model, ICO is committed to the physical proximity of research to clinical activity in order to ensure better results in terms of knowledge transfer. Hence considerable effort has been made in recent years to foster research. According to this, ICO had three research programmes: Epidemiology and prevention of cancer, Translational research and Clinical research.

It is a great pleasure to announce the launching in 2008 of a fourth research program devoted to Basic research, the "Cancer Epigenetics and Biology Program" (PEBC) headed by Manel Esteller, M.D., Ph.D. The research program is supported by the Departament de Sanitat de la Generalitat de Catalunya, Ministerio de Sanidad y Consumo, Ministerio de Ciencia e Innovación and the Catalan Institution for Research and Advanced Studies (ICREA). Current research projects developed at the PEBC include, but are not restricted to, cancer epigenetics, chromatin and stem cell biology, aging, cell cycle and transformation, imprinting and the identification of cancer genes. Representative highimpact publications from the Principal Investigators of PEBC include Nat Genet 37:391-400, 2005; Nat Cell Biol 7:930-2, 2005; Nat Genet 37:1063-71, 2005; Nat Genet 38:566-9, 2006; Cell 125:719-32, 2006; Genes Dev 20:1256-61, 2006; Nat Rev Genet 2007; Nat Genet 39: 279-80, 2007; Nature 450: 440-4, 2007; N Engl J Med 358:1148-59, 2008.

We encourage applications from highly motivated candidates with excellent qualifications in biomedical and life sciences, molecular and cellular biology or biochemistry for the following positions:

- 1 Junior Group Leader (Includes funding for reagents and personnel)
- 1 Research Associate 3 Postdoctoral Researchers
- 3 PhD Students
- 3 Technicians (Including a confocal specialist)
- 1 Scientific Management Officer

The place of employment and work is the Cancer Epigenetics and Biology Program of the Catalan Institute of Oncology, located at the Parc Sanitari Duran i Reynals, Av. Gran Via s/n Km 2,7 08907 L'Hospitalet de Llobregat, Barcelona, Catalonia, Spain. Applicants are required to submit a curriculum vitae in English, a copy of the degree certificate and a letter of interest to Ms. Anna Verges (averges@iconcologia.net) or Dr. Manel Esteller (mesteller@iconcologia.net). Any potential applicant is invited to apply irrespective of age, sex, race, religion or ethnic background. We encourage unsolicited applications from excellent applicants at all times.

For further information, please visit <http://www.pebc.cat>



W160543R



The Institute of Chemical Research of Catalonia (ICIQ) 2008 Offer of PhD Studentships in Chemistry

The **Institute of Chemical Research of Catalonia (ICIQ)** has a number of internationally leading research groups which are active in a diverse range of areas including **catalysis, renewable energies, computational chemistry and supramolecular chemistry**. PhD studentships are currently available for the following projects:

- ✓ Multimolecular Functional Assemblies (*P. Ballester*)
- ✓ Computational Modeling of Homogeneous Catalysis (*C. Bo*)
- ✓ Immobilized Systems with Enzyme-like Behavior (*M. A. Pericàs*)
- ✓ Catalysts for Artificial Photosynthesis (*A. Llobet*)
- ✓ Development of New Gold-Catalyzed Reactions (*A. M. Echavarren*)
- ✓ New Challenges in Catalytic C-H Bond Functionalization (*R. Martín*)
- ✓ Hierarchical Zeolites: Enhanced Utilization of Microporous Crystal by Advances in Material Design (*J. Pérez-Ramírez*)
- ✓ Synthesis of Bowl-Shaped Molecules (*A. M. Echavarren*)
- ✓ Selective Extraction of Actinides from Nuclear Waste (*J. de Mendoza*)
- ✓ Covalent and Supramolecular Approaches to Catalysis (*A. Vidal*)
- ✓ Development of new Carbonylation Catalysts (*P. W. N. M. van Leeuwen*)

The candidates should have (or expect to obtain) a degree in Chemistry, Chemical Engineering, Physics or a related discipline. The studentship provides a stipend comparable to an F.P.U. M.E.C. fellowship rate and covers the tuition fees for the enrollment in the graduate program of the University Rovira i Virgili. The fellowships are open to students of any EU country. For research details visit the website www.iciq.es. For informal discussions, please contact any supervisor using the e-mail address on the website.

Completed application form (available to be downloaded at ICIQ's website), **including the names of two referees, should be sent by e-mail to fellowships08@iciq.es as soon as possible and at the latest by 30th July 2008.**



W161477R

CNIO-CAJA NAVARRA INTERNATIONAL POSTDOCTORAL PROGRAMME

✖✖✖ a project
chosen by
our clients **can** ✖✖✖

CALL FOR APPLICATIONS

Spanish National Cancer Research Centre **cnio**

The **CNIO-Caja Navarra International Postdoctoral Programme 2009** is calling for highly motivated individuals to apply for training within the CNIO's research programmes on **Molecular Oncology, Cancer Cell Biology, Structural Biology and Biocomputing, Molecular Pathology** and **Human Cancer Genetics**.

For details on the individual research groups and recent publications, please see the CNIO's 2007 Scientific Report available at www.cnio.es/ing/sr.

We seek:

- Outstanding young scientists from any country with a Ph.D./M.D. degree obtained no more than two years prior to the date of application.
- Candidates without a Ph.D./M.D. degree may apply provided they obtain their degree before February 2009.
- Candidates must be first author (including joint first authorships) of at least one publication in a high impact factor, peer-reviewed journal at the time of application.

The CNIO offers:

- Excellence in research and a stimulating working environment embracing a multidisciplinary approach to cancer research.
- All scientific activities at the centre are conducted in English.
- State-of-the-art equipment and core technologies.
- A competitive salary of € 32,000 to € 36,000 (gross) per year, depending on experience.
- A comprehensive benefits package as well as in-house English and Spanish language classes.

Secure your interest today by sending an application and you could be part of one of the European cancer research centres of excellence.

The call closes on December 31st, 2008. However, applications will be considered at regular intervals. For further information and to apply on-line visit www.cnio.es/postdoc or contact the Training Office at postdoc@cnio.es.

Centro Nacional de Investigaciones Oncológicas (CNIO)

Melchor Fernández Almagro, 3, 28029 Madrid, Spain

Tel: +34 912 246 900 | Fax: +34 912 246 980 | www.cnio.es



W164196R



**INSTITUTE
FOR RESEARCH
IN BIOMEDICINE**

IRB Barcelona invites applications for new Interdisciplinary Postdoc positions

This new initiative aims to foster collaborations within IRB Barcelona's research areas. Successful candidates will carry out projects involving two of the following programmes:

- Cell and Developmental Biology
- Structural and Computational Biology
- Molecular Medicine
- Chemistry and Molecular Pharmacology
- Oncology

Applicants should have a PhD degree in a life sciences subject and a strong motivation to work in an interdisciplinary environment.

Applications from biologists, chemists, and structural and computational biologists are encouraged. Fellowships will be funded initially for 2 years.


Generalitat de Catalunya

 UNIVERSITAT DE BARCELONA

 Parc Científic
de Barcelona

Deadline for applications: **September 15, 2008**
For more information, see: www.irbbarcelona.org

W160064R



Advancing the frontiers of Light



© photos Luis Montenegro

Unique premises, unique location: Barcelona

ICFO - The Institute of Photonic Sciences, is active in Barcelona, Spain, since 2002. The young institute aims to push the limits of scientific and technological knowledge in key areas of Optical Sciences and Technologies.

ICFO is engaged in research, education and training of MSc and PhD students and post-doctoral researchers.

Research at ICFO is organized in four wide areas:

Nonlinear Optics	Nano Photonics
Quantum Optics	Bio Photonics

ICFO hosts 45 laboratories and nanophotonics fabrication facility in a 9000 sq.m dedicated building based at the Mediterranean Technology Park.

Unique opportunities:

ICFO+, ICFO Young Minds, ICFO Junior Group Leaders, ...

A central part of ICFO's mission is the education of young scientists through their participation in cutting-edge research, and specialized courses and programs, very specially the ICFO+ program.

Junior Group Leaders at ICFO are tenure-track appointments to young scientists who wish to built-up a world-class group to conduct independent research. ICFO provides outstanding start-up packages, which may be supported by private endowments.

www.icfo.es



W163631R



Your research. Our world.

The **Bioteools Group**, a leading Spanish corporation in the biotechnology sector, keeps strengthening its R&D solutions for the molecular biology area by applying its own patented technologies. Therefore Bioteools provides a wide range of products from recombinant enzymes for PCR (**Bioteools B&M Labs, S.A.**), Diagnostic Kits (**Bioteools Diagnostics**) to the latest microarray system (**Proteogene**).

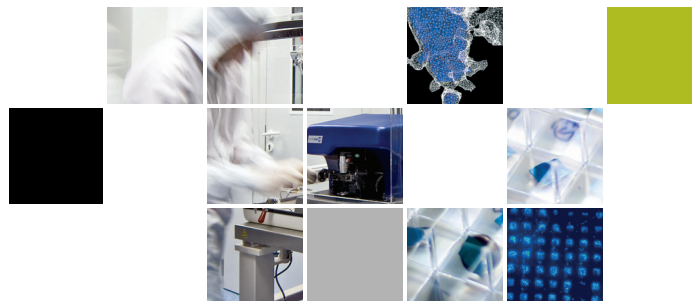
A set of companies committed entirely to the research and development of diagnostic products and solutions.

If you are enthusiastic about advancing in the Life Science world and enjoy great challenges, either in R&D or Sales & Marketing areas, send us an e-mail to careers@bioteools.net and we will be pleased to welcome you on board in the creation of a new future.



www.bioteools.net

W163329R



**Institute for Bioengineering
of Catalonia**

The **Institute for Bioengineering of Catalonia**, located in Barcelona, develops interdisciplinary research of excellence, from basic research to medical applications, in the biomedical engineering and nanomedicine fields.

IBEC focuses its activity on six research programmes: cellular biotechnology; nanobiotechnology; biomechanics and cellular biophysics; biomaterials, implants and tissue engineering; medical signals and instrumentation; and robotics and biomedical imaging.

IBEC is constantly seeking outstanding highly motivated candidates for postdoctoral and PhD positions.

Applicants should send a cover letter and a CV to jobs@ibecharcelona.eu indicating the research programme of their interest.

www.ibecharcelona.eu

R164337W



LOOKING FOR TALENTED AND HIGHLY CREATIVE PEOPLE!

Andalucía is the southernmost region of Spain, the most populated region (+ 8 millions inhabitants), and the second largest one (18% of the national territory). Andalucía also reflects a period of Spain's history in which different cultures coexisted peacefully and each made their greatest contribution in relative harmony for more than 500 years. The impact of this unique commingling still stamps our cities making the dialogue between innovation and tradition stronger than ever.

Recent research advances have given rise to new hope in the treatment and diagnosis of many illnesses that today have no alternative solution. Translational research in the fields of cell therapy and regenerative medicine, genomic medicine, and nanomedicine seem to hold a key role to this riddle. As a result, the Regional Government of Andalucía steers the Andalusian Advanced Therapies Initiative as a priority research objective.

In this sense, Andalucía was one of the first regions in the world that approved legislation to authorize investigation using embryonic stem cells as well as to allow the application of techniques as cell reprogramming, including somatic nuclear transfer therapy. It reflects a strong commitment by the Regional Government to establish an integral strategy, combining various legal, financial, and infrastructural resources in order to create the ideal conditions for the advancement in biomedical research, always keeping in mind a primary goal: to offer the latest therapeutic opportunities available to our fellow citizens, in the framework of a universal health care system.

To achieve these goals, a veritable network of infrastructures for research encompassing three research programs are been developed:

- Andalusian Program for Cell Therapy and Regenerative Medicine
- Andalusian Program for Clinical Genetics and Genomic Medicine
- Andalusian Program for Nanomedicine

Each of the above strategic research programs have a specific thematic research centre.

The first thematic research centre, **CABIMER** (Andalusian Center for Molecular Biology and Regenerative Medicine), was founded in 2006. CABIMER aspires to be a premier focus for research, training and technology transfer in cell therapy and regenerative medicine. Located in Seville, CABIMER occupies a building of over 9,000 m² fitted with cutting edge scientific equipment. Total investment in the centre has exceeded 15 million euros since his foundation. The building has twenty laboratories and different core facilities resources, distributed in three floors, where up to 200 scientists can develop their research activity. The centre also includes GMP (good manufacturing practices) rooms for the preparation of innovative drugs, thereby providing support for clinical trials on cell therapy and regenerative medicine. These

GMP facilities are scattered also within the region with a total number of twelve. Two additional thematic research centres are in the building process with a combined budget of 25 million euros to be ready for work in 2009. These are **GENYO** (Centre for Genomics and Oncological Research) and **BIONAND** (Andalusian Centre for Nanomedicine and Biotechnology), located in Granada and Málaga, respectively. We expect to host around 250 scientific staff to work in pharmacogenomic, molecular oncology, human genetic variability, nanosmart drug delivery, biosensors, and new functional nanoparticles.

To further ensure a translational network, the Andalusian Stem Cell Bank, located in Granada, was the first structure created in Spain specifically for research with stem cells. This centre alongside with the Andalusian Human DNA Bank to be opened in 2009 and the research institutes in close proximity with hospital centres –such as the Institute of Biomedicine of Seville (IBIS) or the Maimonides Institute for Biomedical Research in Córdoba (IMIBIC)- represent a fertile ground for competitive research.

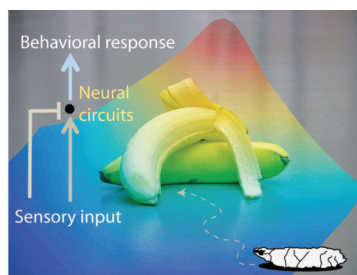
"A network of research centres in a single organization with a universal health care system covering 9 M inhabitants in 41 hospitals...The truth key for your achievement and success in your research career"

If you want to be included in our database, **Andalucía is now actively recruiting** highly-skilled technician, business development staff, pre- and post-doctoral students as well as junior and senior scientists that remain committed to our founding vision: that we can conquer debilitating chronic diseases through research and technology. The success of this effort requires a cohesive team of scientists and clinicians with diverse areas of expertise, but with a shared mission and dedication to the larger goal.

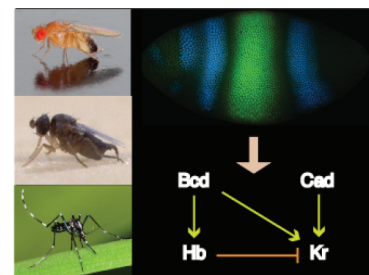


W164633R

Interested Candidates info.celltherapy@juntadeandalucia.es ☎ +34 955 407 131 📠 +34 955 040 457



Laboratory of Sensory Systems and Behaviour



Laboratory of Comparative Analysis of Developmental Systems

The **Centre for Genomic Regulation** (CRG, www.crg.es) is a new research institute located in the Barcelona Biomedical Research Park (www.prbb.org) and associated with the University Pompeu Fabra (UPF). The CRG comprises a Systems Biology unit formally affiliated with the European Molecular Biology Laboratory (EMBL). Research in this unit aims at understanding the integrated function of complex biological systems. Two fly laboratories have recently joined the EMBL-CRG Systems Biology unit and are seeking to recruit motivated researchers.

The **Laboratory of Sensory Systems and Behaviour** led by Matthieu Louis (matthieu.louis@crg.es) is interested in understanding how behavioral decisions emerge from the integration of sensory information. This problem is tackled in the genetically manipulable olfactory system of the fruit fly *Drosophila melanogaster*. A cross-disciplinary strategy involving fly molecular genetics, behavioural analysis, electrophysiology and computational modeling is used to identify, and functionally characterize, the neuronal circuits underpinning larval chemotaxis. Positions are available for one PhD student and a post-doctoral researcher with experience in electrophysiology or biophysics.

The **Laboratory for Comparative Analysis of Developmental Systems** led by Johannes Jaeger (voginho@gmail.com) is involved in a pioneering effort to characterize the developmental and evolutionary dynamics of an entire gene network. We focus on the gap gene system—involved in segment determination during early development—in *Drosophila* and different species of flies, midges and mosquitoes (diptera). To achieve this, we use computational models (gene circuits) as tools to extract the regulatory structure of the gap gene network from quantitative gene expression data. Positions are available for one PhD student and a post-doctoral researcher with experience in either developmental genetics or complex dynamical systems, and a strong motivation and desire for interdisciplinary work.

Candidates should send a full CV, a letter of motivation and the names of min 2 potential referees to blanka.wysocka@crg.es as well as the group leader of interest. All documents should be written in English.

W164642R



Grow Big in a leading company

Like you, we want to continue growing. Our pioneering spirit has driven us to become a consolidated multinational pharmaceutical company in continual international expansion. Our products are sold in over 80 countries including the main markets of Europe, USA and Japan. We have subsidiaries in Europe and Latin America and alliances with the most prestigious pharmaceutical companies worldwide. We focus our research resources in areas related to the treatment of asthma, COPD, psoriasis and rheumatoid arthritis.

We invite you to be part of our projects and reinforce our team! We are constantly attracting talented people for a wide range of positions and areas like R&D (Discovery, Preclinical Development, Pharmaceutical Development and Medical Project Management), Drug Safety, Regulatory, Medical Marketing Affairs and International Marketing, among others.

If you are interested in working with us,
please address to our web site www.almirall.com
or send us an email to rrhh@almirall.es



Solutions with you in mind

www.almirall.com

W164658R



Live, learn and enjoy on the Mediterranean Coast

OFFICIAL POST-GRADUATE STUDIES IN CHEMICAL, ENVIRONMENTAL AND PROCESS ENGINEERING at the ROVIRA I VIRGILI UNIVERSITY (Tarragona, Catalonia, Spain)

Study with a **distinguished** faculty. Expand your mind and horizons with students from over 20 countries. Live in a **multicultural** environment. Make new friends. Enjoy the **beautiful and historic** Catalan environment in Tarragona, UNESCO World Heritage City, only 90 km away from Barcelona "The City of Marvels".

The **excellence** of this international programme is fully recognised by the Catalan and Spanish accreditation boards. We also participate in several international **exchange student programmes** (ERASMUS, MAE-AECI, EU projects, etc.).

We are preferably looking for **recent graduates** seeking an outstanding post-graduate programme focused on **state of the art research and teaching** to participate in our post-graduate studies **fully offered in English**:

Official Master in

- Chemical and Process Engineering
- Environmental Engineering

Official Doctoral Programme in

- Chemical, Environmental and Process Engineering




One-year studentships for Masters and three-year **teaching assistantships** for Doctorate are offered annually for the most qualified applicants. Holders of one-year studentships will be preferably considered later for teaching assistantships. Scholarships are **offered once a year** (May-June) but there is a continuous registration period for candidates with their own support or those who can apply for external funding.

For **more information** and admission conditions, consult our website or contact us by e-mail, telephone or fax.

<http://www.etseq.urv.es/DEQ/Doctorat/web/>

Universitat Rovira i Virgili
Escola Tècnica Superior d'Enginyeria Química (ETSEQ)
Av. Països Catalans, 26, 43007 Tarragona, Catalonia (Spain)
e-mail: secdoc.deq@urv.cat Ph.: +34 977 559720 Fax: +34 977 559699

Organised by:

 School of Chemical Engineering
 Department of Chemical Engineering
 Department of Mechanical Engineering

Sponsored by:  UNIVERSITAT ROVIRA I VIRGILI

 GOVERN DE SPAIN  MINISTERIO DE CIENCIA E INNOVACION  Agencia de Gestió d'Ajuts Universitaris i de Recerca  Programa Mundial DHD

W162664R

IDIBAPS

Institut
D'Investigacions
Biomèdiques
August Pi i Sunyer

IDIBAPS is a leading biomedical research center housed at the University of Barcelona's Faculty of Medicine and the Hospital Clínic, two centenary institutions in the center of Barcelona (Spain).

It offers a wide range of technological facilities, thus providing an important platform for translating basic research into innovation.

At IDIBAPS you will find everything you need to develop your personal and professional potential.

Do you want to join us?

4 POSTDOCTORAL RESEARCH POSITIONS

The 4 Challenging career paths will be developed in these 9 areas:

- Mitochondrial regulation of cell death and steatohepatitis
- Role of glutamatergic system in developing brain damage in chronic hypoxia
- Immunopathogenesis of HIV infection and HIV vaccine development
- Selective gene function modification to promote kidney regeneration
- Mechanisms of B cell dysfunction: Glucolipotoxicity and apoptosis
- Computational Genetics in Digestive Diseases
- Piglet model of ventilator-associated pneumonia
- Proteomics in human tissues of Parkinson's syndromes
- Transcriptional regulation during cell differentiation and oncogenic transformation

Applicants must have received a PhD or MD no longer than four years prior to their application and have a strong background in the chosen area. Written and spoken English is essential. The positions are for three years and include housing facilities.

Deadline for applications: August 14th, 2008

Application forms and more information at: www.idibaps.ub.edu/jobopportunities

W163836R



inibic
instituto de
investigación biomédica
de a coruña



Institute of Biomedical Research of A Coruña (INIBIC)

The INIBIC emerges from the union of Biomedical research groups at the University Hospital Juan Canalejo (1.440 beds), the University and the Juan Canalejo Foundation in A Coruña (Spain). This Institute currently comprises around 3.000 square meters with clinical and basic research laboratories. More than 150 researchers work full time on around 65 projects -which are granted by public agencies, national and international- and clinical trials in collaboration with pharmaceutical companies.

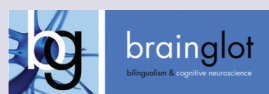
The projects are aligned to the following topics:

- Clinical and Molecular Biology of Rheumatic Diseases
- Cardiomyopathy and Cardiovascular Genetics
- Respiratory diseases: asthma, lung cancer, TB, respiratory Toxicology
- Neuroendocrinology, Obesity
- Rehabilitation in Parkinson, cortex-thalamus interaction
- Cellular Biology on Ageing
- Molecular Biology of Blood Coagulation
- Studies of antibiotic resistance, epidemicity and virulence in *Acinetobacter baumannii*
- Biomedical IT systems, management of digital image and data for Health
- IT for Health, Telemedicine and Biotechnology
- Clinical Epidemiology
- Stem cells, Cell Therapy and Tissue Engineering
- Immunobiology of Transplantation
- DNA and chromosome damage (Toxicogenetics)
- Minimal residual cancer disease, biomarkers and cancer stem cells.
- Health care research by nurses

The INIBIC intends to strengthen its teams by recruiting predoctoral and postdoctoral scientists. Those interested are invited to send a CV to rrhh.fundacion@canalejo.org

For further information:
www.canalejo.org www.udc.es

W164347R



BRAINGLOT PROJECT

Post-doctoral and PhD Positions in Cognitive Neuroscience

Applications are invited for full-time Post-Doctoral and PhD positions in the "BRAINGLOT Project" (<http://www.pcb.ub.es/brainglot>), a Spanish Research Network on "Bilingualism and Cognitive Neuroscience" (Consolider-Ingenio 2010 Scheme, Spanish Ministry of Science and Education).

The project brings together the efforts of six research groups integrating researchers from different scientific disciplines with the common purpose of addressing the phenomenon of bilingualism. The project is conceived with an open and multidisciplinary vocation, as one of its major anchor points places the stress on the mutual influence (both in terms of cognitive and neural processes) between bilingualism and other functions such as auditory perception, multisensory integration, and the executive control of attention. This is an excellent opportunity for professional growth for those interested in the fields of Psychology, Cognitive Neuroscience, Linguistics or related disciplines including Computer Science.

Postdoctoral Positions: Candidates must have a completed PhD and strong background in Developmental Cognitive Neuroscience, Neuroscience, Cognitive Psychology or related field. The positions are funded for 2 years. Competitive salary according to profile and experience.

PhD Positions: We are seeking highly motivated applicants with a background in Cognitive Neuroscience, Life Sciences or Computer Science.

Applications should include a C.V. and the following supporting documentation (letter of interest, publications, awards, letters of recommendation). Send applications to the Project Manager: **Jaume Condal**, jcondal@pcb.ub.es

Deadline: 31th December 2008. Please, use these references in the email subject: Postdoc-BRAINGLOT or PhD- BRAINGLOT.

W164333R

The Centre for Research on Agricultural Genomics (CRAG) CSIC-IRTA-UAB is an association formed by the CSIC (the largest multi-disciplinary research institution in Spain, www.csic.es), IRTA (the agro-alimentary research institute of the Autonomous Government of Catalonia, www.irta.es), and the Universitat Autònoma de Barcelona (www.uab.es). The main goals of the CRAG are to advance in the knowledge of plant and farm animal biology using molecular biology and omics approaches and to supply innovative tools to agriculture in the areas of variety improvement and diagnoses.



The CRAG is presently located in the city of Barcelona and a new building that will host the Centre is under construction in the Campus of the Autonomous University of Barcelona in Bellaterra and it is expected to open in the second half of 2009.

Positions of both scientific, technical and managerial staff will be offered through the web page www.csic-irta.es

The CRAG is funded by a grant from the "Direcció General de Recerca, DIUE, Generalitat de Catalunya".

The CRAG is currently seeking a **PROJECT MANAGER**.

He/She will have a university degree in Sciences ideally in Biology and experience gained in a similar position for at least 2 years.

The candidate will have good knowledge of both English and Spanish, the ability to work both independently and within a team and will be able to adapt to scientific departments in a multicultural organization.

More information in our web address www.csic-irta.es where a letter of application including a full curriculum vitae should be submitted before 31 July 2008.

W164348R



CENTRO DE BIOTECNOLOGÍA Y GENÓMICA DE PLANTAS (CBGP)

The Centro de Biotecnología y Genómica de Plantas (Centre for Plant Biotechnology and Genomics), is a joint venture of Universidad Politécnica de Madrid (UPM) and Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), created with the aim of being a centre of excellence for Plant Biology in Spain. The new facilities of this research centre are located at UPM's Science and Technology Park at Montegancedo, Madrid (Spain).

Research at CBGP is carried by 19 groups organised around four axes: Plant Development, Plant-Microbe Interactions, Abiotic Stress and Functional Genomics.

New permanent positions for research leaders are created periodically.

A new permanent position at CBGP will be called in September 2008 for a scientist with a Ph D degree, with 6-10 years of post-doctoral experience in one of the following areas:

- Bioinformatics
- Genomics
- Systems biology
- Molecular Genetics of Plants
- Molecular Genetics of Plant-Associated Microorganisms
- Metabolomics

In addition, through its supporting institutions, UPM and INIA, CBGP is within the network of Research Centres accepting **temporal contracts for junior scientists** of Spanish and European Programmes such as Ramón y Cajal or Marie Curie.

Post-doctoral positions at CBGP are also available, funded by the supporting institutions and by the research groups.

Further information on CBGP and career opportunities may be found at its web page (www.cbpg.upm.es), or by directly contacting the CBGP Director (Fernando García-Arenal, fernando.garciaarenal@upm.es) or Executive Director (Antonio Molina, antonio.molina@upm.es).

W164502R



The big challenge of the small



The Nanoscience Cooperative Research Center, CIC nanoGUNE Consolider, located in San Sebastián (The Basque Country, Spain), will open the doors of its own building with state-of-the-art facilities for nanoscience research after summer 2008.

www.nanogune.eu

W164478R



Senior and Junior positions as both Scientist and Process Engineer in Nanophotonics

The **Valencia Nanophotonics Technology Center (NTC)** is a research institute at the Technical University of Valencia, Spain, pursuing excellence in knowledge generation, development and integration of innovative photonics technologies and systems.

Both senior scientists and senior process engineers will be responsible for R&D projects on nanophotonic devices for a wide range of applications (ICT, Photovoltaics, Space, Defence, Biophotonic sensors...). The successful candidates will lead (seniors) or contribute (juniors) to their teams using state-of-the-art simulation tools and CMOS-compatible processing equipment in our new 500m2 clean room in close collaboration with DAS Photonics, our highly successful spin-off company, and other industrial and academic partners.

Profile:

- Bachelor's degree in Engineering, Physics or Chemistry (PhD would be advantageous)
- Hands-on approach, motivated and innovative, teamwork orientated.
- Good oral and written communication skills in English. Some Spanish would be helpful

Our offer:


To attract the most dynamic and talented professionals, we are offering a highly attractive compensation and benefits package, including a performance-orientated salary for the senior position commensurate with qualification and experience. The type of contracts available may depend on the candidate's profile and will be awarded following a formal call for applications.

Interested candidates please submit your comprehensive CV in confidence to jobs@ntc.upv.es indicating in the subject line of the email the word "junior" or "senior" and the name of the website where you first saw the advertisement. Particularly for the senior position please state your current / expected salary on your cover letter or CV.

Contact information
Javier Martí, email: jobs@ntc.upv.es, Ref: JOB3n08.
Web: www.ntc.upv.es

July 2008

W164521R



Head of the Cellomics (Cytometry & High Content Analysis) Unit (Ref. HC)

The CNIC (Spanish National Center for Cardiovascular Research) is looking for an experienced and highly motivated specialist in Cytometry and High Content Analysis (HCA) to direct the newly-established Cellomics Unit. The new service head will play a crucial role in defining the medium and long term strategic development of the HCA and Flow Cytometry service, and one of the most important activities will be to instigate and coordinate research programs aimed at advancing the technological potential of HCA. Wide experience in flow cytometry, image acquisition and processing, screening design, and data analysis is requested. The successful candidate will participate in the communication of results and methodologies with peers in the field, and will be responsible for coordinating the contributions of other technical units and research staff to the High Throughput Screening service. Key activities will include the establishment of basic detection methodologies for use in experimental protocols, the design and implementation of screening protocols, the development of software and computing resources for data storage and analysis and the coordination of research projects and technologies in flow cytometry and HCA.


Interested candidates should send their CV with the reference by email only to: rrhh@cnic.es

Details of our employment policy and further information about the CNIC can be obtained at <http://www.cnic.es>

The final date for submission of applications is July 31, 2008

The CNIC supports equal opportunities regardless of gender or nationality. Absolute confidentiality during the selection process is guaranteed.

W164472R



www.cima.es

Around 400 professionals from over 20 countries in Europe, Africa, America and Asia work in CIMA of the University of Navarra. We emphasize translation of basic knowledge to clinical practice.

GROUP LEADER - ATHEROTROMBOSIS

The Division of Cardiovascular Sciences develops its activities in hypertension and atherothrombosis. It bases on the delineation of normal and pathological biological mechanisms and the exploitation of the emerging understanding of these mechanisms to identify biomarkers of disease. Such biomarkers must facilitate the identification of disease subtypes and point the way toward new molecular targets for prevention, diagnosis, and treatment. CIMA invites applications for a tenure-track Group Leader position in Atherothrombosis. The successful candidate will have a PhD or MD and an excellent post-doctoral career.


Contact: Javier Diez, MD, PhD (Head of Division) - jadimar@unav.es

TWO SENIOR SCIENTISTS

- MOLECULAR NEUROPATHOLOGY (PARKINSON & ALZHEIMER)
- MOLECULAR BIOLOGY (ALZHEIMER)

The Division of Neurosciences intends to set up state-of-the-art research programs in Alzheimer's and Parkinson's disease through the study of its genetic, molecular and pathophysiological basis. The laboratories explore biomarkers and new therapeutic approaches ranging from the use of adult stem cells and trophic factors to promote neuronal regeneration and differentiation to clinical research on deep brain stimulation. CIMA invites applications for two senior scientist positions: one expert in molecular neuropathology focused on Parkinson's and Alzheimer's disease and one expert in molecular biology of Alzheimer's disease. Successful candidates will be given the opportunity to lead a research group within their expertise area. Knowledge of Spanish is recommended.

Contact: Julio Artieda, MD, PhD (Head of Division) - jartieda@unav.es



cima
CENTER FOR APPLIED MEDICAL RESEARCH
UNIVERSITY OF NAVARRA
cima@unav.es - Avda Pio XII, 55, E-31008 Pamplona. Spain

W163935R

Commitment to healing, passion for life

I WORKSHOP - DIFHEMAT**From Lymphoid Differentiation to Cancer****Madrid - November 20-21 2008 Application deadline:** September 15, 2008**Venue:** CNIC, Melchor Fernández Almagro, 3, Madrid, Spain**Organisers:**

I. Moreno de Alborán, M.A. Rodríguez-Marcos, A.R. Ramiro, M.L. Toribio

Confirmed speakers:

F.W. Alt Harvard Medical School, USA; **M. Busslinger** Research Institute of Molecular Pathology, Austria;
S. Casola IFOM-IEO, Italy; **F. Dieterlen-Lièvre** Institut d'Embryologie, France; **A.A. Ferrando** Institute for Cancer
 Genetics, Columbia University, USA; **T. Graf** Centre de Regulació Genòmica, Spain; **P. Herbomel** CNRS, Institut Pasteur, France;
E.J. Jenkinson MRC Centre for Immune Regulation, UK; **P.W. Kincade** Oklahoma Medical Research Foundation, USA;
J.P. Manis Harvard Medical School, USA; **I. Moreno de Alborán** Centro Nacional de Biotecnología, Spain;
J. Moscat University of Cincinnati, USA; **A.R. Ramiro** Centro Nacional de Investigaciones Oncológicas, Spain;
B. Rocha INSERM, Institut Necker, France; **M.A. Rodríguez-Marcos** Centro de Biología Molecular, Spain;
F. Sánchez-Madrid Hospital Universitario de La Princesa/Centro Nacional de Investigaciones Cardiovasculares, Spain;
M.L. Toribio Centro de Biología Molecular, Spain; **J.C. Zúñiga-Pflücker** Sunnybrook Research Institute, Canada

For more information and application:
www.cbm.uam.es/difhemat



W164198E

**Inbiomed**

This non-profit research foundation is dedicated to translational stem-cell research and regenerative medicine. Its location at the thriving Technology Park of San Sebastián favors strong interactions with local hospitals and biotech companies. The main goal of Inbiomed is to advance basic knowledge of human stem-cell biology and clinical translation by developing cell-based therapeutic approaches. Inbiomed research is focused on four strategic programs: epithelial, hematopoietic/mesenchymal, neural and cancer stem cells.

Inbiomed harbors an advanced technological platform in cell therapy including the first adult stem cell bank in Europe, Inbiobank. It also hosts a GMP-certified pharmaceutical laboratory for cell therapy, Inbiopharma, and the necessary expertise to develop cell-based products suitable for clinical application. Currently Inbiomed is responsible for the production of bioengineered human skin in a Phase II clinical trial for epidermolysis bullosa. For more information please visit www.inbiomed.org.

W164647R

**GROUP LEADER POSITIONS AT THE IBGM**

The "Instituto de Biología y Genética Molecular" (IBGM), a joint research institute of the University of Valladolid (UVA) and the Spanish Research Council (CSIC) performs basic and applied research in the following areas: molecular basis of cell activation, chemoreception, inflammation, neurodegenerative processes and embryonic development, programmed cell death, and genetic diagnosis of perinatal diseases and cancer (www.ibgm.med.uva.es). We are now seeking to fill two open group leader positions. Selected candidates will be invited for a symposium held at the IBGM. The positions are permanent and associated with Spanish civil servant status and require fluency in Spanish.

To apply please send a CV and a summary of past and future research interests to: director.ibgm@csic.es

Closing date for applications: August 31st 2008

W161583R

www.themobilitynode.cat

Helping researchers' mobility

coordinated by FUNDACIÓ CATALANA
PER A LA
RECERCA
I INNOVACIÓ

W162300R



Institut de Recerca Biomèdica - IRB, Barcelona, Spain
 is seeking

Molecular Biologist - Mouse Mutant Core Facility

You will provide expert technical support for researchers generating various types of gene targeting and transgenic constructs, and will be expected to provide advice (and practical help) in cloning and screening strategies. You will also be required to identify useful novel cloning methods in the literature, and implement them where appropriate.

We require a highly motivated and skilled molecular biologist, capable of creative thinking. The candidates should have previous experience of generating gene targeting constructs, preferably by homologous recombination in bacteria ('recombineering').

Experience in mammalian cell culture and knowledge of written and spoken Spanish would also be an advantage.

Application by CV with covering letter to: sara.martorell@irbbarcelona.org

Application deadline: 1st August 2008

Quote ref: Lab/08/03

W163616R

“Naturejobs.com effortlessly delivered high calibre applicants with the right credentials. Giving the opportunity to quality scientists and giving us the competitive edge”

Erik A. Miljan, PhD

Head of Stem Cell Discovery

ReNeuron Limited, United Kingdom

**Trip Hazard**

Career paths can be precarious. Stride confidently with insider-knowledge from Naturejobs.

naturejobs

Misprint

Buyer beware.

Vonda N. McIntyre

Fluffy was crying.

Like a baby.

My new cat's name is FluffIII, the third in the Fluffy line, but I call him Fluffy. I'm not sure what I'll call FluffIV, if there is one.

The new cat is supposed to be a normal ordinary catprint, a match to Fluffy, the best cat ever, who I raised from a kitten. With a catprint, you get a grown cat. That skips the kitten part, which is too bad, but you also skip housetraining. And neutering.

A couple of selection boxes. Neural path: Use the litterbox. A deletion: Testicles.

Correct?

Print.

I never bother with the rest of the form. The non-allergenic Hairless option. Hairless cats are the next ugliest thing to naked mole rats anyway.

And who would tick off 'Round Pupils'? Cats should have cat's eyes, slit-pupilled, glow-in-the-dark.

Opushun advertises available Polydactyly. A cat with opposable thumbs? Forget it. Thumbs are fine for helper animals like dogs or ponies, but with a cat you'd have to add Anxious to Please, a high-performance package (translation: extra cost). They claim it works, but I doubt it. Millennia of domestication didn't produce Anxious to Please in cats.

And if it did, you wouldn't exactly have a cat, would you?

So I chose the standards, ignored the rest, and snapped the order.

A week later — it takes that long to grow enough cells for the printer cartridges — Opushun snapped me back to pick up FluffIII.

The clerk opened the carrier. The Fluffys have long soft grey fur and gold eyes. Their paws look silver in some light. You can change all that, but if I did, I wouldn't exactly have Fluffy, would I?

"He looks perfect! Thank you." I reached into the cat carrier and let him sniff my fingers and rub against my hand. I skritchd his ears. I got a little teary. One reason Fluffys are such great cats is they're friendly. They're not anxious to please, but they are, naturally, friendly.

I thumbed for FluffIII and took him home.

But now he was crying.

I jumped up, startled. My work connection broke. I go to another place when I work so I like it to be quiet where I really am.

"Fluffy, nice kitty, what's the matter?"

He looked up from his kitty bed.

"You never talk to me."

He had a perfect Oxbridge accent, that plummy high-class British way of talking.

"I was just talking to you. I talked to you all the way home."

He licked his silver paw and rubbed the tears from his eyelashes. "You call that a conversation? 'Nice kitty, kitty?'"

"I didn't expect you to understand. I didn't expect you to talk back."

"I am not talking back. I am trying to hold a civilized conversation."

"I didn't think cats had civilized conversations," I said.

"Hah," he said. "Or, to be more precise, LOL."

He jumped into my lap, I cradled him, and he managed to look dignified with his paws splayed all which-way. I petted his belly. The Fluffys have very soft fur.

"We'll have to converse later,"

I said. "Work calls." I put him back in the kitty bed. Being upset, I expended time and effort to reform the connection.

Fluffy strolled into the room, his furry tail high.

"When's dinner?"

The work connection collapsed.

Before Fluffy could protest, I shoved him into the cat carrier.

"There's been a mistake," I said to the clerk at the Opushun kiosk. I hauled Fluffy out. He glared at me without a word.

"I'm so sorry!" the clerk said. "At Opushun we're anxious to please."

"Can you fix him?"

She frowned, puzzled. "Surely it's already fixed?"

"No, yes, I mean, he came already fixed, that isn't the problem. The problem is..." I glanced down at Fluffy. "Tell her the problem."

"That is the problem," Fluffy said. "I'm fixed. Can you fix *that*?"

The clerk pretended he wasn't talking.

"Sir, you may exchange any model that's defective."

"Defective!" Fluffy said.

The clerk blinked. For a second I thought she had eyes like Fluffy's but then I blinked and she looked normal again. People's parents don't alterate them outside

human traits, usually. Maybe she was wearing mood contacts.

"You see?" I said.

"I'm terribly sorry, sir, I don't."

"He's *talking*."

"Our new development," she said. "It's so popular that we made it standard. At Opushun, we're —"

"— Anxious to please," I said. I hadn't even noticed available Conversation so I hadn't opted-out, which would have been an extra charge because they'd have to upgrade the standard by downgrading the upgrade.

"You see?" said Fluffy.

"Can you change him so he doesn't talk?"

Fluffy said something unprintable, which was startling in that accent. The clerk wasn't pretending he couldn't talk — she was ignoring him. That seemed rude. If you're going to create something that talks, you should listen to it.

"I'm so sorry, sir. Alterating a printout is un-cost-effective. But you may certainly make an exchange."

She tried to return Fluffy to the carrier. His claws scrabbled at the flaps. Probably he wished I'd gone for available Polydactyly.

"Wait!" he said, dignity abandoned. "What happens to me?"

"They'll find you a good home," I said. "With somebody who appreciates your finer qualities."

"Prove it!"

"I'm sorry, sir," the clerk said. "It's imprinted on you. Another client would not be ... pleased."

Fluffy leaped, flung himself against my chest, and nuzzled my neck, his nose damp and cool. He didn't say anything.

I sighed. "Will you promise not to talk to me when I'm trying to work?"

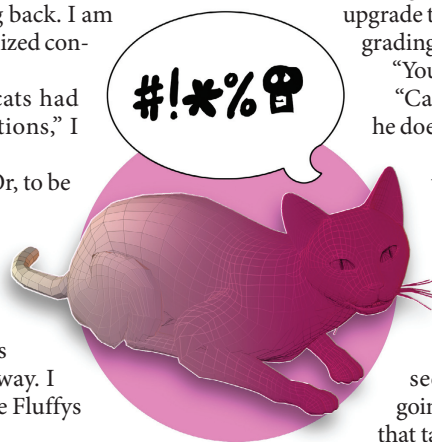
"I'll ... try." He sounded more doubtful than anxious.

I took him home.

He isn't exactly Fluffy.

But he's close enough.

Vonda N. McIntyre once worked in a department that now genetically engineers finger-webs for rainy Seattle's residents. She believes a story isn't worth telling if it isn't worth exaggerating. Her online home is www.vondanmcintyre.com.



JACEY